DISRUPTIVE TELECOMS

Enable. Innovate. Transform

WORLD REPORT July 2026

REINVENTING TELECOMS WITH AI

From connected to intelligent: The next frontier for mission-critical enterprises

J.P. Singh
Nokia

Why Enterprise Networks Must Move Beyond Monitoring

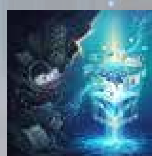
Nadeem Akhtar
HFCL

The Space Economy's Biggest Revenue Engine Isn't Launch. It's Connectivity

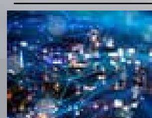
Lt. Gen. AK Bhatt (Retd.)
ISpA

From Connectivity to Trust: How eSIM is Powering the Next Generation of Connected Vehicles

Rahul Tandon
IDEMIA Secure Transactions



Why AI-Driven Communications Are Eating the Collaboration Market



Rewriting the RAN Playbook

An Initiative by TelecomDrive.com

Telecom Drive

THE EDUCATION PROJECT

A unique platform by TelecomDrive.com
to put spotlight on innovators – and next
generation innovation driven by global universities,
technology institutes and industry ecosystem

TelecomDrive.com

From the Editor

Zia Askari
Editor, TelecomDrive.com

Beyond the Hype:
Why Agentic AI Will Define the Next Decade of Telecom

As we look toward the future, Artificial Intelligence is poised to fundamentally reshape the telecommunications industry. The conversation has shifted from a question of if AI will impact telecoms to how fast, how safely, and at what cost. The industry is at an inflection point where AI is transitioning from a novel tool for operational efficiency to the core engine of business strategy and network infrastructure.

The Emergence of Agentic and Autonomous Networks

The most significant trend on the horizon is the move from simple automation to "agentic AI"—autonomous software entities capable of reasoning, planning, and executing complex workflows without constant human intervention. This evolution is driving the industry toward fully autonomous network operations. According to Omdia research, 41% of Communication Service Providers (CSPs) identify network management as the most significant area of impact for agentic AI, signaling a fundamental shift toward self-healing, self-configuring networks.

This is not a distant vision; major vendors like Amdocs and Google Cloud are already collaborating to deliver production-ready, agentic AI solutions for telco contact centers, bridging complex BSS/OSS systems with advanced conversational AI to automate billing, plan changes, and customer care journeys.

Real-World Returns on Investment

The future of AI in telecoms is increasingly defined by tangible business outcomes. Industry leaders are already reporting significant gains. Deutsche Telekom has achieved 20-30% efficiency improvements from AI micro-agents in network management, while Nokia reports a 30% performance boost with 99% less human effort. As the network becomes more complex with 5G and the coming of 6G, this level of automation will become indispensable. The integration of AI is also crucial for managing energy consumption, with Rakuten, for example, achieving 17% energy savings by deploying AI micro-agents in its Open RAN environment.

Data is the Critical Battleground

Despite the promise, significant challenges remain. The success of advanced AI depends on clean, unified, and real-time data—a major hurdle for many telcos operating with siloed legacy systems. The future will see CSPs creating "ontological AI," where agent behavior is governed by a structural truth of the business, ensuring that AI agents do not develop conflicting interpretations of critical services. Vendors like Wavelo are already helping CSPs wrap their legacy systems with event-driven platforms to feed AI applications with the real-time data they require without the need for costly and risky rip-and-replace migrations.

Evolving Business Models and Enterprise Value

Looking beyond operations, AI is set to redefine the business model. Telcos are transitioning from pure-play connectivity providers to strategic "techcos" and enablers of enterprise digital transformation. They are developing "AI-as-a-Service" offerings and creating domain-specific proprietary LLMs. This allows them to move beyond simple connectivity and offer differentiated services like predictive analytics and security, creating new, recurring revenue streams. Frost & Sullivan predicts that over the next five years, telcos that invest in AI centers of excellence and modern data infrastructure will shift from experimentation to scaled innovation, driving high-value B2B solutions.

The future is one where intelligence is the new infrastructure. AI is not just a tool for optimization but the foundational layer for a more reliable, efficient, and innovative telecommunications ecosystem.

CONTENT

Reinventing Telecoms with AI : From Network Pipes to Autonomous Intelligence Engines

For decades, the telecommunications industry operated on a predictable playbook: build networks, sell connectivity, and compete primarily on coverage and price. The arrival of generative AI and agentic automation is shattering that model. Telecom operators are no longer merely providers of "dumb pipes" through which digital services flow—they are becoming AI-native technology companies capable of delivering autonomous network operations, hyper-personalized customer exp eriences, and entirely new revenue streams built on intelligence rather than bandwidth.

Network Automation: The Journey from Reactive to Self-Healing Networks

The most transformative impact of AI is occurring deep within the network itself. Traditional network operations have been reactive—engineers respond to alarms, troubleshoot outages, and manually reconfigure systems after failures occur. This model is becoming unsustainable as 5G and future 6G networks grow in complexity, heterogeneity, and scale.

Predictive Analytics and Autonomous Healing

AI is enabling a fundamental shift from reactive monitoring to predictive, self-healing networks. Research has demonstrated AI-augmented hybrid models that integrate dynamic network slicing with self-healing capabilities, achieving fault detection accuracy of 95.2% and traffic demand prediction accuracy of 92.8%. These systems leverage supervised and unsupervised machine learning algorithms to anticipate failures before they impact service quality.

The results are dramatic: self-healing capabilities have reduced fault recovery times to an average of 300 milliseconds, improving system uptime to 99.8% compared to traditional approaches. This represents a 50% reduction in latency and a 28.6% improvement in resource efficiency.

Looking toward 6G, researchers are developing AI-driven self-healing network models that incorporate deep learning-based anomaly detection, reinforcement learning-based fault prediction, and automated reparation policies supported by Software-Defined Networking and edge intelligence. By leveraging real-time telemetry, spatio-temporal traffic patterns, and multi-domain contextual data, these models achieve early fault prognostication and intelligent root-cause analysis—essential for the ultra-reliable, zero-touch networks of the future.

Agentic AI in Network Operations

The evolution extends beyond predictive analytics to fully autonomous operations. O2 Telefónica in Germany, in partnership with Tech Mahindra and NVIDIA, has launched a Large Telco Model (LTM) to power its "Operations of the Future" program. The LTM enables agentic capabilities across the operations lifecycle—automated root-cause analysis, field dispatch optimization, and service experience management—all through an end-to-end, dynamic view of critical services.

Similarly, Google Cloud and Nokia have expanded their partnership to introduce a multi-agent framework powered by Gemini models. The system features six specialized AI agents handling event analysis, anomaly detection, performance interpretation, remediation recommendations, dashboard generation, and cross-agent orchestration. Nokia estimates the AI-driven architecture could reduce troubleshooting times from hours to minutes while reducing false alarms and unnecessary escalations.

This represents what Nokia describes as "glass box autonomy"—a model that keeps human engineers involved in decision-making while allowing AI agents to provide confidence-based recommendations. Operators retain final approval over critical actions, though low-risk scenarios can be automated through closed-loop processes.

From Scripts to Intent-Driven Automation

The challenge of telecom operations has long been the reliance on domain experts and ad-hoc scripts for network reconfiguration—a process that is labor-intensive, error-prone, and fragile. Generative AI is offering a fundamentally different approach. Researchers have introduced intent-driven autonomous reconfiguration frameworks that integrate generative AI with hierarchical multi-agent architectures. At their core, a chief agent transforms natural-language intents into dependency-aware execution graphs, delegating atomic tasks to specialized agents responsible for IP management, network control, cloud orchestration, and service management.

Customer Experience: From Chatbots to AI-Native Intelligence

The customer-facing transformation is equally profound. For years, telecom customer service has been defined by rigidity—static menus, scripted chatbots, and DIY apps designed primarily for one-way interaction. Generative AI is replacing this with something genuinely conversational.

Conversational AI and the End of Menu Fatigue

Instead of searching through plan catalogues or waiting on hold, users can now describe their needs in plain language: "I'm traveling next week and need data in Europe but not the US." Intelligent agents can instantly compare plans, summarize differences, and even adjust offers through natural dialogue. In demonstrations, customers have "negotiated" with AI, asking it to swap data for extra minutes or adjust pricing to fit their preferences. The result is a dynamic, two-way conversation ending with an active eSIM download and a sense of genuine personalization.

Maxis in Malaysia recently introduced Miya—a generative AI-powered conversational assistant built with Google Cloud's Customer Engagement Suite. Miya handles bill management, roaming recommendations, and contract reviews through natural conversation rather than menu navigation.

AI Concierges and Zero-Agent Models

Circles, a global technology company powering digital telcos, has launched an AI concierge built on OpenAI's API platform. The platform features CareX, powered by a proprietary multi-agent architecture that orchestrates specialized agents to handle end-to-end tasks including network diagnostics and billing—with a 95% resolution rate. It currently resolves 85% of global queries without human intervention, delivering instant support at scale.

Complementing CareX is Circles Xplore IQ, an AI-powered monetization engine that predicts customer needs before they're expressed, matches those needs with personalized offers, and autonomously completes transactions. During its initial phase, the system delivered an ARPU uplift of 22% for Circles. Life Singapore while reducing churn by 9%. The company's goal is a fully autonomous telco maximizing profitability and customer satisfaction.

Some digital-first operators are already exploring "zero-agent" models where customer issues are handled entirely by AI from ticket creation to resolution. This shift will unfold in phases, beginning with hybrid setups where AI handles routine inquiries while humans supervise complex cases.

AI as a Revenue Driver, Not Just a Cost Saver

The commercial implications extend beyond operational efficiency. Operators are increasingly bundling premium AI services into mobile plans. Perplexity Pro has been partnered with Deutsche Telekom, Singtel, Sunrise Switzerland, Optus, and Telkomsel, giving over 250 million subscribers access to AI-powered search. ChatGPT Plus has similarly partnered with Orange, SK Telecom, and O2 Telefonica Germany.

Consumer reception is cautiously optimistic. CCS Insight research found that 26% of UK consumers would sign up for free access to premium AI services, with another 32% saying their decision depends on which tool is offered. Crucially, 84% of 25- to 34-year-olds expressed interest—suggesting that as generational adoption of AI tools grows, so will the premium service market.

The key differentiator is that AI enables operators to blend efficiency with empathy—doubling down on automation while also being able to "understand" and relate to users' needs in context. As one industry observer noted: "Customers feel understood, not targeted, because the suggestions are relevant and well-timed".

Energy Efficiency: AI as the Network's Green Engine

As 5G networks expand and 6G approaches, energy consumption has become an urgent concern. Base stations account for 73% to 77% of total network energy usage, with a typical 5G macro base station consuming approximately 3.3 to 9 kW. AI is proving instrumental in addressing this challenge.

AI-Native RAN Energy Optimization

Research has demonstrated AI-driven energy savings within Open RAN architectures through what is known as the Energy SavingrApp (ES-rApp). This AI solution integrates traffic-aware prediction models and closed-loop automation to dynamically optimize network operations—enabling adaptive cell deactivation,

intelligent transmit-power adjustment, and proactive resource allocation based on predicted traffic demands.

The results are significant and validated on real testbeds: ES-rApp achieves up to 19.5% energy savings without degrading quality of service. This represents real-world evidence that AI-native energy optimization can be a scalable, practical solution for green RAN management.

AI-Driven BSS/OSS: The New Operating System for Telecom

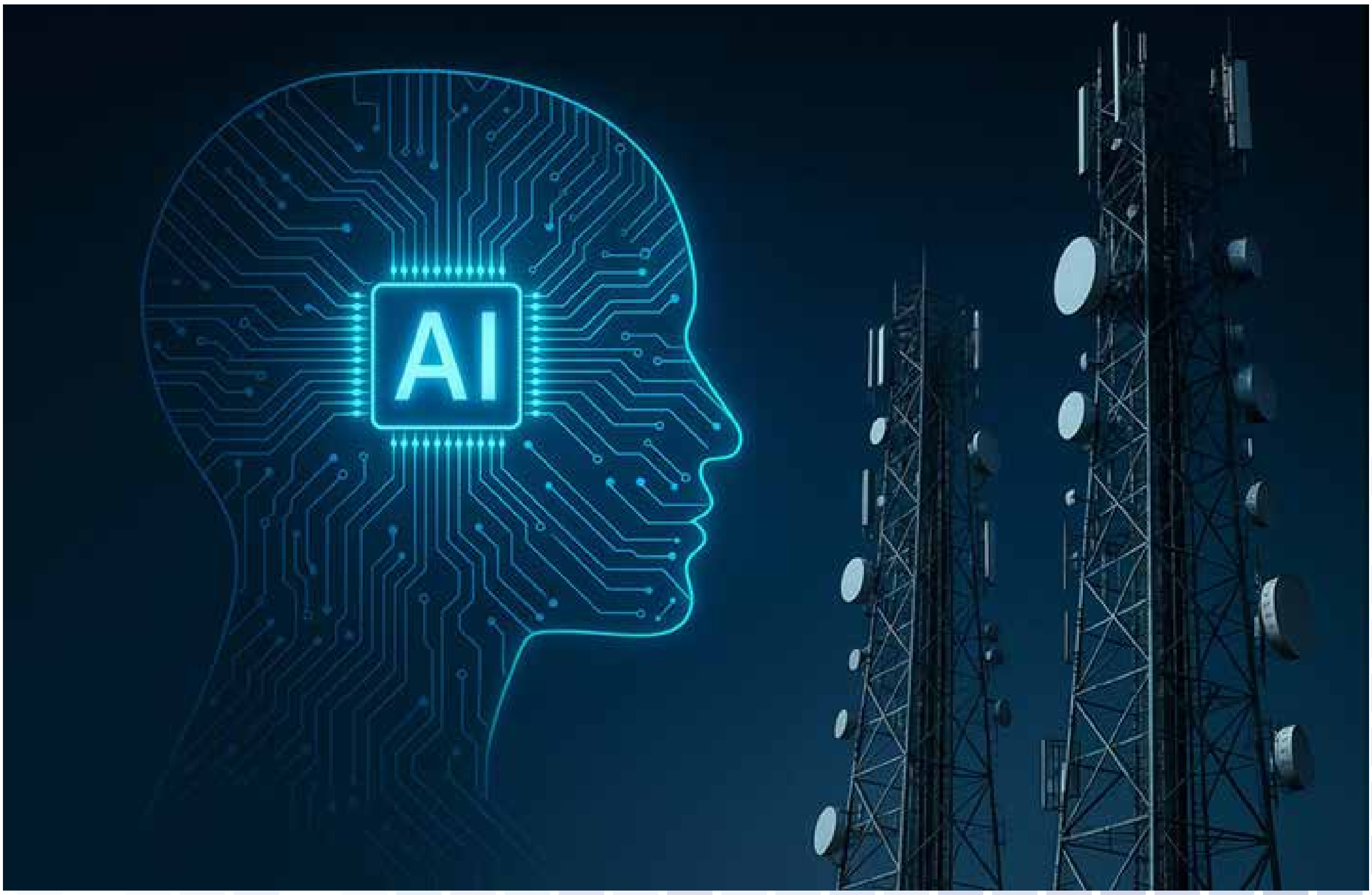
The transformation of Business Support Systems (BSS) and Operations Support Systems (OSS) represents perhaps the most consequential shift in telecom IT strategy today. For decades, these systems operated as separate, siloed domains—BSS handling customer care, billing, and revenue management; OSS managing network inventory, fault resolution, and performance. Generative AI and agentic architectures are dissolving these boundaries, creating what industry leaders describe as a new "operating system" for telecommunications.

From Siloed Automation to Agentic Orchestration

The first wave of AI in BSS/OSS focused on point automation—chatbots for customer service, predictive analytics for network faults. The current wave is fundamentally different. Agentic AI—systems that understand goals, make decisions, access data, and execute across tools without direct human intervention—is enabling coordinated, multi-step process execution across previously disconnected systems.

This shift is visible across the vendor landscape. Amdocs has introduced aOS, an agentic operating system that operates on top of any BSS/OSS stack, embedding intelligence directly into telecom operations. The platform features a Cognitive Core with pre-built, telco-specific agent libraries, enabling CSPs to design and run agentic processes across customer care, service orchestration, monetization, and network operations. The architecture supports multiple large language models and maintains telco-grade security and observability—critical for mission-critical systems.

Ericsson has adopted a complementary approach with its OSS/BSS Business Value Pathways, launched in partnership with AWS. The framework provides pre-designed, integrated operational combinations to ensure autonomous network transformation outcomes are delivered. Rather than simply adding AI to legacy automation, Ericsson's pathways offer holistic transformation across three high-impact areas: accelerating monetization, enabling experience, and improving efficiency. The collaboration with AWS uses generative AI to enrich service orchestration, turning static workflows into intent-driven



execution where systems can autonomously resolve issues and keep services running.

The Challenge: Legacy Systems and Data Quality

Despite the momentum, significant barriers remain. In some cases, operators will still need to modernize or extend underlying BSS and OSS systems. The telecom industry is saddled with legacy systems, siloed infrastructure, and imperfect data quality that can impact the efficiency of AI agents . An Omdia survey found that 46% of telecom operators expect GenAI to impact their BSS and OSS business within the next two years, but the path to realization requires careful planning .

The Human Factor: Moving to Supervisory Roles

The shift to autonomous operations requires a parallel shift in workforce roles. Comarch emphasizes a “Human-in-the-Loop” methodology where AI provides deep operational insights and handles routine tasks, but key decisions regarding infrastructure changes are supervised and authorized by experts. This ensures a safe, gradual transition to greater autonomy, building trust in AI systems while allowing operators to retain responsibility and control. As Omdia analysts note, CSPs must transform employees from task executors to agent supervisors, defining policies that agents must follow and handling exceptions agents cannot resolve.

Standards and Openness: Guardrails for Transformation

To protect against vendor lock-in and ensure interoperability, industry analysts recommend CSPs adhere to TM Forum’s Open Digital Architecture and Open API standards. Initiatives like TM Forum’s OmniBOSS project demonstrate how AI-powered, agent-orchestrated systems can sit above existing OSS/BSS tools, normalize data using TM Forum frameworks, and identify gaps between actual operations and best-practice processes . The result is consistent operational discipline across multi-vendor environments.

Beyond Cell Sleeping

The BeGREEN project, supported by the European Commission, has consolidated final results on AI/ML-assisted energy efficiency for 5G and 6G-ready RANs. The research quantifies the energy footprint of the AI/ML plane itself—training, serving, storage, and orchestration—against the network-side savings it drives, demonstrating that lightweight models achieve favorable trade-offs between size, accuracy, and inference rate.

The project evaluated five AI/ML-assisted methods to enhance energy efficiency, including intelligent cache management for platform power reduction, carrier on/off classifiers that trigger sleep/activation with guardrails, and multi-objective coordination that maintains QoS while cutting total energy across static, heterogeneous, and dynamic scenarios.

Looking to 6G, researchers are exploring Fog Intelligent RAN

architectures that internally integrate AI to reduce power consumption and improve energy efficiency. Studies show that F-IRAN maintains better efficiency levels than conventional F-RAN, particularly as the number of network devices increases.

The Business Model Transformation

Perhaps the most significant shift is how AI is changing telecom business models. As Frost & Sullivan analysts note, generative AI offers telcos the opportunity to move "beyond selling connectivity and network services, evolving into strategic enablers of enterprise digital transformation".

New Revenue Streams

Telcos are exploring multiple AI-driven revenue models: AI-as-a-Service: Embedding NLP, computer vision, and GenAI into new B2B offerings through subscription-based pricing models

In-house Proprietary LLMs: Creating domain-specific telecom LLMs that facilitate AI-based network optimization, intelligent edge solutions, and predictive analytics

Vertical and Horizontal AI Solutions: Developing turnkey AI solutions that integrate seamlessly into enterprise workflows, including white-label services, chatbots, and customized foundational models

The Software-Defined Future

Realtransformation requires fundamental changes to network architecture. And operators cannot afford for 6G to be a major capex spend without new business models. The linchpin is the move to software-defined infrastructure, which allows operators to update capabilities in near real-time and forms the foundation for AI-native 6G.

More than 20 operators worldwide have announced AI data centers—smaller than hyperscale builds but strategically placed to support regional needs like data privacy and mission-critical workloads. Telcos are always kind of a trusted entity by the governments and by the enterprise and research for data privacy, for mission-critical equipment, for regulations.

Conclusion: The AI-Native Telco

The reinvention of telecom through AI is not a distant vision—it is happening now. Networks are becoming self-healing and predictive, customer experiences are becoming conversational and personalized, and energy efficiency is being optimized through intelligent orchestration.

The operators who succeed in this transformation will be those who embrace AI not as an efficiency tool but as a fundamental reimagining of their business. As Frost & Sullivan concludes, "AI places telcos at an inflection point: moving from pure-play connectivity and network services to AI-powered technology companies".

The AI-native telco will be defined by autonomous operations, predictive intelligence, and customer experiences that feel less like navigating a utility and more like having a personal assistant. The transition requires investment, cultural change, and a willingness to move beyond traditional boundaries—but the opportunity, as the industry is discovering, is transformative.

From connected to intelligent: The next frontier for mission-critical enterprises

By J.P Singh, Market Head India - Mission Critical Enterprise, Government& ROI at Nokia

Over the last decade, mission-critical enterprises around the world have made a significant investment on connectivity by linking assets, people and systems across some of the most demanding environments, from remote mines to public safety operations. That bet has largely paid off.

But for leadership teams, a harder question is now taking shape: having connected the enterprise, what is the return on making it intelligent? This is no longer a technology question. It is a business one about risk, resilience, and the cost of decisions made too slowly or on incomplete information.

Connectivity was always the foundation, not the outcome. Its value is only realised when a connected operation can also sense, decide, and act in real time,particularly in moments where there is no room for error, and where a delay of even seconds carries real financial or safety consequences.

The hidden cost of fragmentation

Across segments, enterprises have invested in strong, purpose-built systems, whether for fleet tracking, worker safety, situational awareness,or network performance, and each performs well in isolation.

The issue is what happens between these systems, not within them. Built by different vendors, on different data models, with no common language, they rarely work together as one operational picture. The result is a familiar and costly pattern: warning signs that exist in the data but are missed in practice, decisions that wait on manual correlation, and risk that surfaces only after it has become an incident.

Consider a mining operation where a truck flags a mechanical issue. The signal is there. But without an automated way to connect it to road-condition data, the underlying cause stays hidden until the asset fails, production stops, and the cost shows up in downtime and repair, not just in a maintenance log.

This is the strategic gap facing leadership today: enterprises are data-rich but often lack the insights needed to act efficiently. The infrastructure to be intelligent already exists. What's missing is the layer that makes it actionable.

Why this matters at the leadership level

For the company's leadership, the implications go beyond any single incident. Fragmented operations mean:

- Slower response, because insight depends on people manually joining the dots across tools
- Hidden risk, because problems are visible in isolated data sets long before anyone connects them
- Constraints on scale, because every new site, vehicle, or system adds complexity

None of this is a failure of any one system. It is the natural outcome of solving problems one tool at a time, in an environment that now demands something more integrated.

The way forward: a unified intelligence layer

The answer is not simply another point solution, nor a multi-year integration project layered on top of existing complexity. It's a unified intelligence layer; one that sits across the entire operation, understands context across systems, and enables action at the exact point where a decision needs to be made, not after the fact in a control room.

For this to work in the environments these industries actually operate in, it has to perform where centralised, cloud-first architectures typically struggle: remote sites with limited connectivity, harsh operating conditions, and situations where real-time decisions are measured in milliseconds, not minutes.

What this delivers, in practice

The specifics vary by industry, but the underlying capability is the same: connecting fragmented signals into one intelligent operational scenario and enabling action at the point where it matters most.

In a mining operation, for example, this could mean closing the gap between environmental conditions and asset health, catching stress on equipment before it turns into unplanned downtime, protecting both production continuity and capital assets. In a defense or field-operations context, it could mean communication that adapts automatically across whichever network is strongest at that moment, whether 4G/5G, Wi-Fi or even satellite, so that critical information reaches decision-makers without interruption, regardless of which infrastructure happens to be available. In an emergency response scenario, it could mean every unit arriving at an incident contributing immediately to a shared operational picture, rather than waiting for a dedicated command vehicle to arrive, so situational awareness scales in step with the response itself.

Different operating environments, but the same underlying pattern: fewer surprises, faster and better-informed decisions, and operations that scale without a proportional rise in complexity or risk. Regardless of the sector, any enterprise running high-stakes, distributed operations faces some version of this same gap.

The foundation this requires

Delivering on this vision requires more than smarter software layered on top of what already exists. It requires a foundation designed for it from the outset, built on three principles:

1. Resilient communications that use the best available network path and adapt instantly as conditions change
2. Real-time decision-making at the point of action, without dependence on distant cloud infrastructure
3. Compute at the edge, so intelligence runs where the work actually happens, not where connectivity happens to be strongest

Together, these principles turn a connected enterprise into a genuinely intelligent one, able to see more, respond faster, and act decisively at scale.

A strategic choice, not just a technical one

Closing the gap between connected and intelligent operations is not something any single organisation does alone. It requires the right technology partner; one that understands both the digital architecture and the physical realities of mining sites, defense operations, and emergency response.

For mission-critical enterprises operating in high-stakes, high-consequence environments, this represents a genuine strategic opportunity: to move beyond incremental efficiency gains and build operations that are demonstrably safer, more resilient, and better positioned to scale.

The shift from connected to intelligent is already underway across the industries that can least afford to get it wrong. The organizations that make this shift first will be the ones best positioned to capitalize on the opportunities AI creates. Moving toward a unified intelligence layer provides the foundation for that transition, helping enterprises turn ambition into operational reality and transform connectivity into a source of strategic value.

Nokia's network expertise and AI-ready architecture can help enterprises design and realise this foundation, turning ambition into a working reality, so they don't just operate more efficiently, but stay ready for whatever the next disruption looks like.

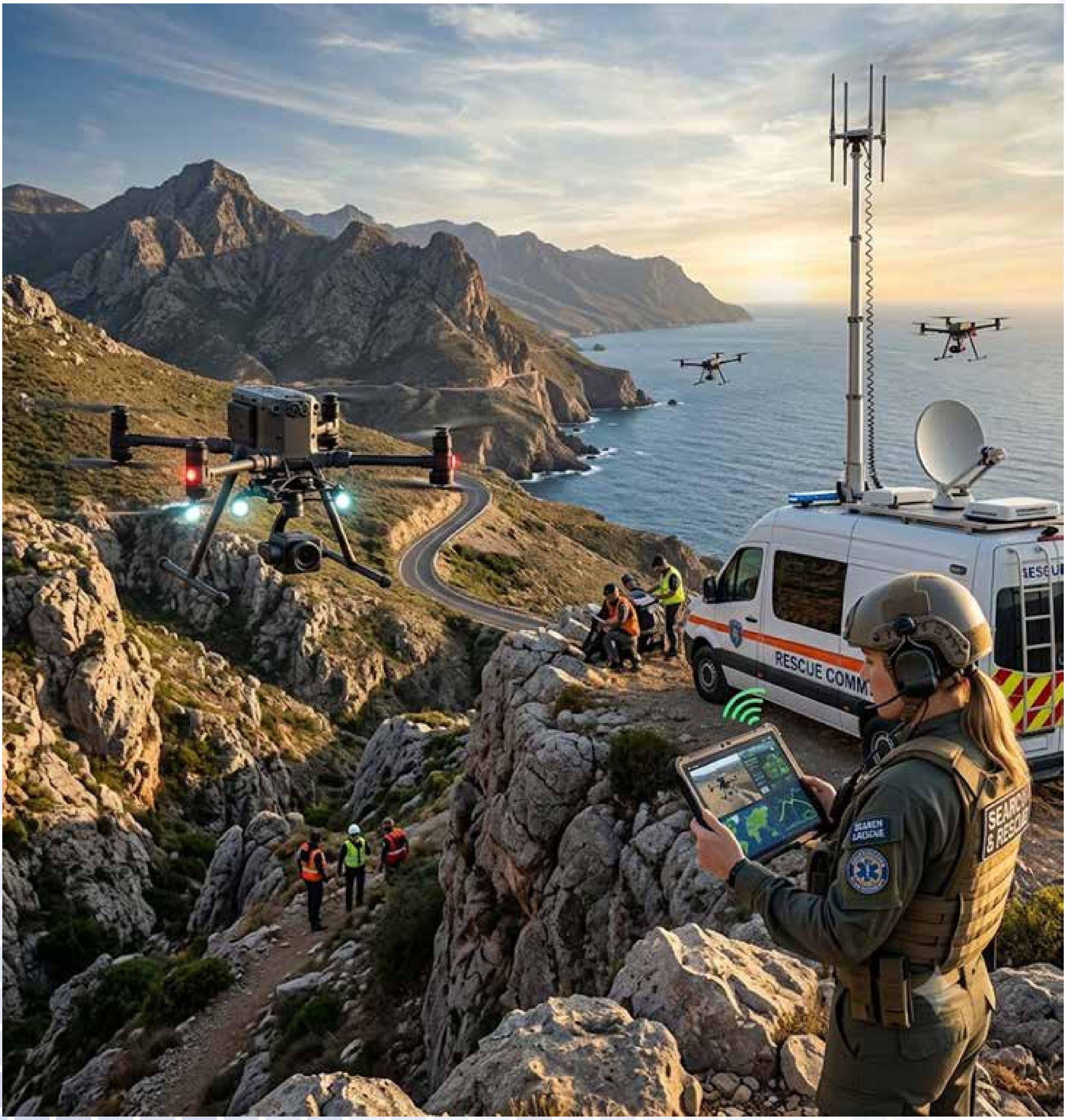
J.P. Singh

J.P. Singh leads the strategy, customer engagement and execution for mission-critical communications and public-sector solutions for Nokia across the Indian market. He has over 20 years of experience in the telecom industry and has held senior leadership roles, including Vice President at Alcatel-Lucent before joining Nokia. Over the course of his career, he has worked across sales, project management and business leadership, building a strong track record with large-scale network service providers and enterprises.

Beyond Voice:

Drones, Resilience, and the Broadband Revolution in Critical Communications

The broadband revolution transforming public safety and utilities



For decades, mission-critical communications have been synonymous with dedicated land mobile radio (LMR) technologies such as TETRA, P25, and DMR. These systems were purpose-built to deliver highly resilient, low-latency group communications for public safety and other critical users. Broadband, by contrast, was originally engineered for consumer data.

That distinction is now dissolving. Mission Critical Services (MCX), standardized by 3GPP, represent the first time that broadband technology has been formally specified to meet the stringent requirements of mission-critical group communications. With 3GPP networks nearing the point where they can fully replace legacy LMR systems, the industry is witnessing a transformation that spans from hybrid voice interoperability to AI-driven disaster resilience and drone-enabled situational awareness.

This article examines the three pillars of this transformation: the migration from LMR to MCX over 5G, the resilience strategies ensuring network uptime during crises, and the integration of drones and sensors that is redefining first responder capabilities.

Mission-Critical Services (MCX) over 5G: Replacing LMR with Broadband

The transition from narrowband LMR to broadband MCX is not a simple technology upgrade—it is a fundamental reimagining of how critical communications are delivered. The 3GPP MCX framework, encompassing Mission-Critical Push-to-Talk (MCPTT), Mission-Critical Video (MCVideo), and Mission-Critical Data (MCData), has matured significantly over recent releases. What began as a standards ambition is now a deployable ecosystem supported by network vendors, device manufacturers, and operators worldwide.

Why the Transition Has Been Measured

Observers sometimes ask why MCX has not displaced LMR more quickly. The answer lies in the magnitude of what is being attempted. Parity with LMR is the baseline requirement. Coverage, resilience, direct mode capability, security, device ruggedness, battery endurance, and operational simplicity all matter deeply. Agencies cannot accept regression in any of these areas.

Due to their proven reliability and critical role in public safety, LMR networks cannot be replaced immediately. The transition from LMR to broadband MCX must follow a phased, carefully managed approach. Past implementations have shown that public safety organizations benefit most from progressive adoption of broadband communications. A hybrid LMR–MCX configuration, particularly for voice communications, represents an effective starting point.

The Interworking Function: Enabling Gradual Migration

The key to this gradual migration is the Interworking Function (IWF), a standardized approach that enables LMR networks to interoperate with 4G/5G-based MCX solutions. A new white paper from TCCA, authored by the organization's IWF Working Group, explores how this integration strengthens communication capabilities, achieves seamless interoperability, and enables a successful adoption of broadband technologies.

The IWF enables seamless voice interoperability between LMR and MCX, data service integration, hybrid network support, multimedia feature enhancements for LMR users, standards-based interoperability, and unified control room integration. The document presents TETRA as a representative case but confirms that the principles can be applied to other narrowband solutions, including APCO P25.

The Migration Timeline

The narrowband-to-broadband transition is unfolding at different paces across the globe. Western and Northern European countries, including the United Kingdom, France, Finland, and Sweden, are already moving ahead with plans to migrate all public protection and disaster relief users from TETRA and Tetrapol systems to nationwide mission-critical 3GPP networks between 2028 and 2031. South Korea is an outlier, having carried out its transition much earlier due to the previous lack of a national-scale digital LMR network.

In the United States, many APCO P25 systems are not expected to be decommissioned until the late 2030s, although some agencies—particularly those whose LMR networks are reaching end-of-life or have poor coverage—are beginning to fully transition to MCPTT services over broadband networks. Romania's TETRA network will continue to operate in parallel with the country's new 3GPP-based PPDR broadband network until 2035.

Global Momentum

Despite these measured timelines, the scale of current investments is substantial. Annual investments in public safety LTE/5G infrastructure and devices reached approximately $5.4 billion in 2025 and are on track to surpass $9.2 billion by the end of 2028, reflecting a compound annual growth rate of around 19 percent. One of the largest projects to emerge from secrecy in 2025 is Saudi Arabia's $8.7 billion mission-critical broadband network for the Kingdom's defense, law enforcement, and intelligence agencies.

National programs extend from

high-profile initiatives such as the United States' FirstNet, South Korea's Safe-Net, Great Britain's ESN, France's RRF, Sweden's SWEN, and Finland's VIRVE 2 to programs across Asia-Pacific, the Middle East, and the Americas.

Resilience and Redundancy: Ensuring Uptime During Disaster and Cyberattack

The promise of broadband MCX extends far beyond faster data speeds. Mission-critical users tolerate no downtime. Communications must perform in routine and crisis conditions, scale across agencies, and defend against evolving threats. That means rigorous engineering around redundancy, prioritization, encryption, and monitoring, while validating behavior under stress and at scale.

3GPP's Resilience Toolkit

3GPP has developed a comprehensive set of features to ensure network resilience for mission-critical users. These include QoS, Priority &Preemption (QPP), which ensures that critical communications take precedence over consumer traffic during network congestion; Isolated Operation for Public Safety (IOPS), which allows networks to continue functioning even when backhaul to the core network is lost; and High-Power User Equipment (HPUE), which extends coverage in challenging environments.

Emerging 5G capabilities further enhance resilience: 5G MBS/5MBS (Multicast-Broadcast Services) for MCX services in high-density environments, 5G NR sidelink for off-network communications, Vehicle-Mounted Relays (VMRs), Mobile gNB With Wireless Access Backhauling (MWAB), and Non-Terrestrial Network (NTN) integration for satellite backup.

AI-Driven Resilience Frameworks

The integration of artificial intelligence is pushing resilience capabilities beyond traditional redundancy. A new AI-Enhanced Sovereign Emergency Communication Network (AI-SECN) framework uses AI, edge computing, and graph-based resilience metrics to ensure independence and resilience during disasters. It introduces a Composite Resilience Score combining path diversity, latency, and localization accuracy. Simulations show the framework increases path diversity by 35%, reduces latency by 22%, and improves overall resilience by 45%.

Hybrid and Satellite Integration

The concept of hybrid networks is central to resilience. Organizations are increasingly combining terrestrial and satellite-based communication to ensure continuity. The KRISNet research project in Germany is developing a highly resilient and sovereign communication infrastructure for digitalized power grids and other critical infrastructures. The project combines terrestrial networks (fiber and mobile) with a sovereign satellite network in low Earth orbit, enabling automatic real-time switching between available networks during disruptions.

FirstNet's macro coverage layer in the United States is already complemented by a growing number of indoor small cells—currently at 14,000 units—supporting operation in Band 14/n14 (700 MHz) spectrum. Britain's ESN, Sweden's SWEN, and Finland's VIRVE 2 programs will also involve large-scale rollouts of in-building coverage solutions.

Drones and First Responders: Real-Time Video and Sensor Integration

The evolution to mission-critical

communications extends beyond transforming field operations—it is paving the way for innovation. With MCX systems, critical communications are becoming true technology hubs, capable of aggregating in real time IoT sensors, CCTV systems, drones, autonomous robots, and mobile cameras. The use of video streams and IoT data opens the door to use cases enhanced by real-time analytics and automated alerts.

Drone as First Responder (DFR)

Drone as First Responder programs are redefining situational awareness by delivering secure, low-latency video from aerial units to command centers and field teams. Traditional response tools can leave agencies reacting with limited information, particularly in critical situations where every second counts. DFR programs address this challenge by transmitting live video directly from the scene before first responders arrive.

Using bonded IP video solutions, agencies can maintain consistent, low-latency connections even when networks are strained or experience disruptions, ensuring uninterrupted visibility across all operational levels. The integration supports scalable, cost-effective deployments that fit within existing communication frameworks, allowing dispatchers, supervisors, and field units to share a unified operational picture.

Automated Emergency Drone Deployment

German research is pushing this concept further. The ADELE project ('Automated drone use from the control center') is developing a system that enables drones to automatically launch as soon as an emergency call is received—even before the first emergency vehicle sets off. The aim is to transmit live optical and thermal images to the control center over the 5G network before emergency services arrive, helping to more effectively plan and execute deployments.

Building on this work, the AI4Relief project investigates how an advanced, AI-supported drone system combined with mixed-reality technologies can improve firefighters' situational awareness. The autonomous drones are integrated into existing control center systems, enabling direct communication with emergency services. On top of fire detection, the system optimizes resource management and dynamically plans deployment routes, based on a deep-learning model that predicts the spread of wildfires, taking into account weather data, topography, and vegetation.

Airborne Reconnaissance Advantages

Satellites can be directed precisely over affected regions, but when dealing with forest fires and other natural disasters—at least in easily accessible areas in or near inhabited regions—satellites generally cannot match the response speed of airborne reconnaissance. Airborne cameras achieve far higher ground resolution (down to a few centimeters), making them ideal for detecting smaller fires and even embers that are difficult to spot from the ground.

Modular aerial camera systems provide high-resolution optical and thermal images that can precisely locate and measure fire sources while documenting their spread. Other tools being researched include intelligent drone and robot swarms that in the future will work autonomously and as teams to investigate forest fires and other acute crisis areas.

The Integration Challenge

The integration of drones, video feeds, and sensor data is increasingly gaining significance in national public safety broadband programs, alongside the inclusion of rapidly deployable network assets, direct-to-device connectivity from satellites, and in-building coverage for emergency communications. The convergence of 5G NR sidelink-equipped prototype terminals for device-to-device communications and multi-hop relay networking—being made available for field trials by defense and public safety agencies between 2026 and 2027—will further expand the capabilities of drone-based operations.

Conclusion: The Coexistence Era

A realistic assessment suggests that coexistence will define the next decade. LMR networks represent significant national investments and continue to provide unmatched narrowband voice resilience. Broadband MCX, meanwhile, delivers data-rich capability and operational flexibility.

Rather than a binary replacement narrative, the industry is embracing integration. Interworking gateways, converged devices, and shared control room environments are allowing agencies to adopt broadband at their own pace while safeguarding continuity.

The trajectory is now clear. Standards maturity, ecosystem depth, field validation, and 5G advancement are converging. What once appeared as isolated puzzle pieces—spectrum policy, device capability, QoS enforcement, interworking, cybersecurity—are increasingly fitting together into a coherent framework. For public safety and other critical communications users, the future will not be defined by voice alone. It will be shaped by integrated, data-driven operations delivered over resilient broadband platforms engineered to mission-critical standards.

Why Enterprise Networks Must Move Beyond Monitoring

By Nadeem Akhtar, Senior Vice President – Product Line Management, HFCL

Most of the excitement around AI in enterprise networks today centres on its most visible form: assistants that create tickets, chatbots that handle routine queries, and generative tools that summarise an incident after it has already happened. These help, but they are not the part that changes how a network is actually run.

The more important shift is happening underneath, in network operations. For most of the past two decades, network management has worked the same way. A system raises an alert, an administrator interprets it, decides what it means, and takes action. Dashboards improved and management software matured, but the workflow remained largely reactive and manual. That model served the industry well when networks were smaller,less distributed and more predictable. The volume of alarms and notifications generated has grown well beyond what reactive, manual workflows were designed to handle.

This is the real meaning behind the phrase "reinventing telecoms with AI." It is less about new interfaces and more about closing the gap between what network reports and what operations teams can realistically act upon.

The reactive model no longer scales

Consider what enterprise networks have become. The move to cloud, the dramatic increase in the number of connected devices, hybrid work models, and a constantly evolving mix of application traffic have made network more dynamic and business critical than ever.

For many IT and network teams, daily operations involve navigating a continuous stream of alerts, managing configurations across multiple sites, and addressing issues only after users begin to complain. Root causes are often reconstructed manually from logs and datasets that were never designed to tell a clear story.

An even deeper issue is that traditional metrics can be misleading.

Quality of Service (QoS) metrics such as throughput, latency, jitter, and packet loss describe what the network is doing, not what the user is experiencing. A network can remain within every predefined performance benchmark while a video conference freezes, a business application slows down, or a user struggles to connect.

Aggregate dashboards can also conceal the experience of individual users whose connectivity has quietly degraded. As a result, organisations often optimise for uptime and healthy-looking dashboards while the experience the business actually depends upon remains unmeasured.

While the network complexity is growing, the IT team is shrinking. This makes the challenge even more difficult to solve without adopting a fundamentally different approach.

Closing the loop

The shift now underway is from monitoring toward more autonomous operation.The objective is a network that continuously observes itself, identifies issues and their impact, takes corrective action where appropriate, and verifies whether those actions achieved the intended result. In effect, the network operates as a closed-loop system, with people supervising outcomes and handling decisions that genuinely require judgment rather than manually triaging every alert.

It is worth being precise about what a credible version of this requires, because "self healing" is an easy phrase to overstate. It works best as a layered intelligence in which autonomy is earned rather than assumed. A foundational layer maintains real-time awareness of every device and client and, importantly, measures experience rather than throughput. An analytical layer converts this stream of signals into actionable insights, identifies root causes, and prioritises incidents according to user impact rather than technical severity.

This means a site-wide authentication failure receives greater attention than a single disconnected device because its business impact is significantly higher. Only then does a further layer act by optimising radio performance, steering clients before they encounter congestion, and automatically correcting well-understood fault conditions.

The discipline lies in that final layer. Autonomous correction should always be governed by confidence. A high-confidence diagnosis can proceed automatically. A medium-confidence recommendation may require human approval. A low-confidence finding should simply be highlighted for investigation.

Every action is logged with the reasoning that triggered it, and the outcome is measured. If user experience does not improve within a predefined period, the system should roll back the change and notify an administrator. The network is empowered to act, but never blindly, and it remains accountable for the result.

The other aspect of this shift is AI-Ops which aims to translate user intent to automated workflows. For instance, when a customer wants to perform a WIDS audit, it should not require clicking across multiple forms in the dashboard, followed by manual analysis of the results. AI-Ops can take the intent/goal (e.g. scan the RF environment for threats), execute the WIDS scan, generate an e-mail summary report and send it to the CIO along with a link to detailed report.

This is the thinking behind the way we have built IO Canvas, our AI powered platform for unified management of wired and wireless networks. At its core is a Quality of Experience engine that translates hundreds of live variables, including signal, roaming behaviour, authentication, DNS and DHCP health, and application performance, into a single, continuously updated measure of what users are actually experiencing. Around it sits a layered architecture that moves from real time visibility to root cause analysis to predictive optimisation, so that channel changes happen without dropping sessions, clients that linger on a weaker connection are steered to a better one, and recurring faults are resolved before a ticket is raised. The point is not the feature list. It is that AI is the foundation rather than an add-on layer.

Why the business changes, not just the workflow

When the loop closes, the economics of running a network change.A relatively small team can manage a larger and more distributed network estate because monitoring, correlation, and first-line resolution are increasingly handled by the system itself. This allows experienced professionals to focus on architecture, optimisation, security, and strategic initiatives rather than repetitive troubleshooting.

The optimisation target also begins to shift from uptime toward experience. That distinction matters because user experience is ultimately what enterprise customers buy and what they are increasingly likely to expect within service-level agreements.

Security also becomes harder to treat as a separate console. When intrusion detection and access control share the same experience aware view, an attack such as a deauthentication flood can be understood not only as a security event but as a measurable effect on users and prioritised accordingly.

Advanced operations are also no longer reserved for the largest enterprises and telecom operators. When the same intelligence can be deployed through lightweight on-premises infrastructure or managed cloud environments, midsize businesses and branch offices gain access to capabilities that previously required dedicated network operations centres.

Much of the practical disruption over the next few years is likely to come from this democratisation of intelligence rather than from the largest operators alone.

The Real Change

It is tempting to frame AI in enterprise networks as a productivity upgrade, the same work done faster and more efficiently. That interpretation understates the transformation. The more meaningful change is structural: a move from networks that report their status toward networks that take more responsibility for their outcomes, and from teams that spend their days reacting to teams that act with better information because the system has already done much of the diagnostic work.

Generative AI will continue to produce the demonstrations that capture attention.The more enduring transformation is taking place within the operations layer, where networks are learning to convert their own data into decisions.That is the direction worth building toward, and increasingly the one worth deploying today.

Nadeem Akhtar

With more than 20 years of experience in telecommunications, Nadeem has worked extensively across Wi-Fi, 4G, and Fixed Wireless Access technologies. He leads HFCL's Wireless and Switching R&D team and is a recognized expert in wireless standards, network innovation, and large-scale connectivity solutions. A frequent speaker at industry events, he holds a PhD in Mobile and Wireless Communication from the University of Surrey and has authored multiple research publications and patents in the field of wireless communications.

The Satcom Gold Rush: How LEO Constellations Are Reshaping Global Connectivity

The rise of LEO constellations and the commercialization of satellite-terrestrial convergence

The satellite communications industry is experiencing a transformation as profound as any in its history. What was once the domain of a few established players with massive geostationary satellites is now a bustling arena where low Earth orbit constellations, direct-to-device connectivity, and hybrid network architectures are reshaping the connectivity landscape. With projections suggesting the D2D market could reach $30 billion by 2035 and cumulative LEO investments exceeding $10 billion by the end of 2026, the opportunity is both immense and urgent.

Direct-to-Device: The Race to Eliminate Cellular Dead Zones

The direct-to-device segment has moved from PowerPoint presentations to commercial reality with remarkable speed. By the end of 2026, the industry expects over 1,000 D2D-capable satellites in orbit, with spending on D2D satellite capacity reaching $6 to $8 billion. What makes this particularly significant is that D2D connectivity is now reaching standard, unmodified smartphones—a development that fundamentally changes the addressable market.

The Key Players and Their Strategies

Starlink has established itself as the dominant constellation, with 9,400 satellites in orbit and a growing number equipped with D2D capabilities.Its vertically integrated model—spanning satellite manufacturing, launch services, and network operations—provides advantages that extend far beyond first-mover status, including end-to-end supply chain control and flexibility over pricing models.

AST SpaceMobile is pursuing a different approach through mobile network operator partnerships via its BlueWalker platform, aiming to deliver mobile-grade connectivity directly to handsets through a planned constellation of 60 satellites. Major operators including AT&T, Verizon, and Bell have committed to delivering D2D services, with commercial beta launches expected in the first half of 2026.

The competitive landscape also features Iridium, Viasat, Lynk, and Sateliot carving out roles in specific regions or specialized applications. In Europe, OQ Technology and Telefónica Germany are conducting the first live demonstration of two-way satellite communications to unmodified smartphones using terrestrial mobile spectrum. China has moved particularly decisively, with all three state-owned telcos receiving regulatory approval for satellite-based mobile services to consumers.

The Monetization Challenge

Despite the momentum, monetization remains a significant hurdle. Consumer willingness to pay is demonstrable—monthly price points of $15-20 are well below legacy satellite-phone tariffs —but the economics remain challenging. Per gigabyte, satellite bandwidth costs 10 to 20 times as much as 4G or 5G. This firmly places current D2D services in the complementary bucket: emergency messaging, basic voice, and low-bandwidth data in areas without terrestrial coverage.

However, the cost per bit of satellite data is reportedly falling by 20-30 percent annually, driven by reusable launch vehicles, mass-produced flat-panel antennas, and laser inter-satellite links. As costs continue to decline and speeds improve—AST SpaceMobile is targeting 120 Mbps directly to smartphones—the line between complement and competitor will thin considerably.

Enterprise and Government Applications

For the enterprise and government segments, the D2D opportunity extends far beyond consumer messaging. Disaster recovery, defense communications, maritime and aviation links, mining, oil and gas, and IoT in agriculture all represent use cases that terrestrial networks struggle to address. The Indian Navy has signed multi-year satcom contracts worth over ₹4,000 crore, while aviation Wi-Fi is projected to become a Rupee Foradian1,500 crore market by 2028. For government and defense applications specifically, D2D satellite connectivity provides critical fallback capabilities when terrestrial networks fail during hurricanes, wildfires, and floods.

Hybrid Networks: Seamless Roaming Between Terrestrial and Non-Terrestrial Networks

If D2D represents the consumer-facing opportunity, hybrid networks represent the architectural foundation that will make satellite-terrestrial convergence truly seamless. The industry consensus has shifted decisively: the future of connectivity is not terrestrial or satellite alone, but hybrid networks that combine both into a single orchestrated system.

The Architecture of Convergence

Traditional satellite and terrestrial networks evolved as separate environments, each with their own core, authentication systems, and management processes. 5G Non-Terrestrial Networks (NTN) change this requirement entirely. To support provider roaming, mobility, service continuity, and enterprise-grade performance, satellite networks must integrate directly with 5G cores and operate as part of a common architecture.

The industry has developed two primary pathways to achieve this integration :

Interworking Gateway Approach: This allows satellite operators using existing non-3GPP infrastructure to access the 5G Network through a Satellite Interworking Gateway (SIG). Operators can continue using established satellite waveforms like DVB-S2X while tapping into 5G Core efficiencies, protecting existing investments while enabling new revenue opportunities.

Native 5G NR NTN Architecture: This approach uses satellite-optimized gNodeB and 3GPP's 5G NR waveform to connect 5G UEs directly to the network. This is the model telcos prefer because it aligns directly with their standards-based RAN architecture and simplifies roaming.

Market Readiness

The appetite for hybrid networks is substantial. A global survey of 350 telecom decision-makers across seven markets found that 95% describe satcom as essential or important to achieving seamless global connectivity, and 88% are planning to increase investment in satellite technologies over the next five years. Importantly, 22% named 5G expansion as their biggest technology focus, with 19% already identifying 5G NTN as a future growth opportunity.

Barriers and Bridges

Cost (52%), complexity (49%), and regulatory challenges (46%) remain the primary barriers to adoption. However, hybrid architectures address these concerns by allowing new capabilities to be introduced incrementally, creating a converged environment where satellite and terrestrial connectivity can be managed as part of a single network rather than separate systems.

The momentum is reflected in recent demonstrations of interoperability between satellite ground systems and 5G cores, validating that non-3GPP satellite networks can appear fully 3GPP compliant while operators continue using established satellite-specific waveforms. This capability enables satellite network operators to extend coverage across new geographic regions and orbital paths seamlessly.

Regulatory Landscape: Spectrum Sharing and Licensing for New Players

The regulatory environment is perhaps the most complex piece of the satcom puzzle, with implications that extend far beyond simple licensing.

The FCC's Spectrum Sharing Evolution

In the United States, the Federal Communications Commission has taken a significant step by modernizing the spectrum sharing framework for geostationary (GSO) and non-geostationary (NGSO) systems. The previous approach relied on Equivalent Power Flux Density (EPFD) limits developed in the late-1990s, which the FCC determined were overprotecting GSO systems and preventing American households from receiving the fastest space-based broadband available.

The new rule replaces the EPFD framework with modern, performance-based GSO protection criteria, extends the framework for good-faith coordination, and allows NGSO and GSO operators to bargain for appropriate interference protections through voluntary private agreements. Technical backstops protect GSO systems when coordination has not been reached. The FCC estimates the rule will generate $1.6 billion to $19.9 billion in economic benefits.

The Coverage-Permission Gap

A critical regulatory distinction that has emerged is the gap between technical coverage and legal permission. A satellite footprint may make connectivity technically available within a state, but lawful service depends on separate permissions for spectrum coordination, station licensing, terminal authorization, market entry, cybersecurity, and numerous other regulatory layers.

This distinction is particularly relevant for satellite broadband mega-constellations like Starlink. Technical coverage does not equal authorized service. National regulators control market access, earth-station operation, terminal licensing, and equipment authorization. The challenge is compounded by the "legal dispersion"—the fact that satellite broadband governance spans multiple authorities, instruments, and requirements.

The Indian Regulatory Debate

India offers a case study in the complexities of satellite regulation. The industry has raised concerns about a "fundamental structural defect" in the proposed Satellite Communication Network authorization framework. The debate centers on whether satellite communication should be authorized as a service under the Telecommunications Act or treated primarily as a network.

The proposed model envisages a network entity that would provide satellite capacity to authorized telecom service providers rather than serving end-users directly. Critics argue this separates network control and spectrum from accountability to end users, potentially blurring spectrum responsibility, raising costs, and delaying connectivity in underserved regions.

The Department of Telecommunications has clarified that a network-authorization entity could seek administrative spectrum assignment where its proposed use falls within the First Schedule's spectrum categories. Critics counter that infrastructure sharing could be enabled separately while keeping spectrum and end-user accountability with the authorized service provider.

The ultimate regulatory choice will determine whether satellite communications develop as a separate ecosystem or integrate seamlessly with terrestrial networks, with significant implications for investment certainty and competitive dynamics.

European Sovereignty and Spectrum Strategy

In Europe, concerns about digital sovereignty are driving a distinct approach. The OQ Technology and Telefónica Germany demonstration of two-way satellite communications using terrestrial mobile spectrum reflects an effort to build sovereign space-based mobile infrastructure. The goal is to develop open standards, licensed spectrum, and existing smartphones to complement terrestrial mobile networks while preserving operator control and spectrum sovereignty.

What Future Holds

The satellite communications opportunity is now unequivocally commercial rather than theoretical. The D2D market is on track for mainstream adoption by 2026, with major operators across multiple regions launching services. Hybrid network architectures provide the technical foundation for terrestrial-satellite convergence, with the market demonstrating clear readiness to invest. Regulatory frameworks are evolving to accommodate new players and business models, though significant complexity remains.

The next twelve to eighteen months will be pivotal. For telecom operators, the choice is not whether to embrace satellite connectivity but how—whether as partners integrating D2D into their customer experience, or as potential disruptees watching global players capture enterprise and premium segments.

For regulators, the challenge is to balance investment attraction and competition with spectrum accountability and consumer protection. For the industry as a whole, the opportunity is substantial: a future where connectivity is truly seamless, universal, and resilient—where the distinction between terrestrial and satellite networks becomes invisible to the user.

The Space Economy's Biggest Revenue Engine Isn't Launch. It's Connectivity

By Lt. Gen. AK Bhatt (Retd.), Director General, Indian Space Association (ISpA)

When people think about the global space economy, they often picture spectacular rocket launches, reusable launch vehicles, lunar missions or ambitious plans for deep space exploration. These achievements have transformed access to space and dramatically lowered the cost of reaching orbit. Yet they tell only half the story.

The real economic value of space is not created at launch. It begins once satellites are in orbit and information starts flowing. According to Novaspace, the global space economy reached US$626 billion in 2025, with space-enabled applications accounting for the largest share of the market. This underscores a fundamental shift in value creation: the future of the space economy will be shaped not by how many assets are placed in orbit, but by how effectively those assets are transformed into always-on digital infrastructure powering industries and economies on Earth.

This marks one of the most significant shifts in the evolution of the global space industry. Satellite communications, once viewed as a specialised capability serving broadcasting, defence and remote connectivity, has evolved into a foundational layer of global digital infrastructure. The convergence of Low Earth Orbit (LEO) constellations, software-defined satellites, virtualised ground infrastructure and multi-orbit architectures is transforming satellite communications from a standalone service into an integral extension of terrestrial telecom networks.

Connectivity as Monetisation Layer of the Space Economy

Historically, the commercial space value chain was relatively linear. Satellites were designed, launched and operated to deliver specific services. Today, that equation has fundamentally changed. Launch creates assets, but connectivity creates recurring economic value.

Every connected aircraft, vessel, remote mine, industrial facility, logistics corridor or disaster response network depends not simply on satellites in orbit, but on resilient, always-available communications. As launch costs continue to decline and satellite capacity expands, value is steadily moving downstream from infrastructure ownership towards managed services, enterprise applications and digital platforms.

This is why satellite communications has become one of the strongest contributors to the global space economy. Today, satellite communications and Earth Observation together account for around 80% of direct commercial space revenues, underlining how connectivity has become one of the principal engines of commercial value creation across the global space ecosystem. It is no longer just another space-enabled service. It is the digital foundation upon which new industries are being built.

Where the Real Downstream Economy Begins

Perhaps the clearest evidence of this transformation lies in how satellite communications unlocks the value of other space technologies.

Earth observation imagery has limited commercial utility if it cannot be delivered fast enough to support time-sensitive decisions. Positioning, Navigation and Timing (PNT) becomes significantly more valuable when integrated with continuous communications. Satellite-enabled IoT allows industries to monitor infrastructure, supply chains and critical assets across geographies where terrestrial networks remain unavailable.

The result is an expanding ecosystem of downstream applications, from precision agriculture and smart logistics to autonomous transport, digital twins, predictive maintenance and AI-powered geospatial analytics. In each case, satellite communications acts as the digital transport layer connecting data, applications and users in near real time. The downstream impact extends well beyond the space sector itself. Industry estimates suggest that every dollar invested in upstream space infrastructure can generate approximately US$2.5 in downstream economic activity, as data is transformed into software, digital services and enterprise applications that create sustained economic value.

When Space Becomes Part of the Telecom Network

This evolution is also reshaping the telecommunications industry itself.

For decades, terrestrial and satellite networks developed as separate ecosystems. That distinction is rapidly disappearing. The standardisation of Non-Terrestrial Networks (NTN) through successive 3GPP releases, coupled with advances in direct-to-device connectivity, software-defined payloads and multi-orbit architectures, is creating a future where users move seamlessly between terrestrial and satellite infrastructure without even recognising the transition.

Tomorrow's networks will not distinguish between terrestrial and non-terrestrial infrastructure. They will simply identify the most efficient path to deliver connectivity. Satellite communications is therefore no longer an alternative to terrestrial networks. It is becoming an integral component of future telecom architecture, extending coverage, enhancing resilience and supporting new digital services across sectors.

The Next Disruption Is an Intelligent Network in Orbit

The next phase of disruption will not be driven merely by deploying more satellites. This evolution is also being enabled by a dramatic shift in industry economics. The cost of satellite capacity has fallen from well over US$100 per Mbps per month in the traditional GEO era to nearly US$1-3 per Mbps on modern LEO networks, fundamentally changing the commercial viability of satellite-enabled services.

Optical inter-satellite links, onboard edge computing, AI-enabled traffic orchestration and integrated 6G architectures are transforming satellites from transparent relay systems into intelligent network nodes capable of processing, routing and managing data closer to where it is generated. In the years ahead, the convergence of space, air and ground networks will redefine connectivity itself, enabling seamless digital experiences irrespective of geography.

India's Opportunity Lies Beyond Launch

For India, this transformation presents a defining opportunity. The country's ambitions should extend beyond launching satellites or manufacturing spacecraft. The larger prize lies in building downstream services powered by satellite communications, integrating space-based connectivity into next-generation telecom networks and creating innovative applications across agriculture, logistics, healthcare, mobility, disaster management and national security.

As India strengthens both its digital economy and private space ecosystem, satellite communications can become the bridge connecting technological capability with economic growth.

Connectivity Will Define the Next Space Economy

The first chapter of the commercial space age was about making access to space affordable. The next chapter is about making space indispensable to the global digital economy.

Satellite communications is no longer simply another application within the space sector. It is the digital infrastructure that enables downstream innovation, accelerates telecom convergence and unlocks entirely new economic opportunities. In the years ahead, the biggest winners in the global space economy will not necessarily be those who launch the most satellites, but those who create the greatest value from the connectivity those satellites deliver.

Lt. Gen. A.K. Bhatt (Retd.)

Lt. Gen. A.K. Bhatt (Retd.) is currently the Director General of the Indian Space Association (ISpA), an apex industry body created to be the collective voice of the Indian Private Space industry. Among his many appointments, he has been Principal Staff Officer (PSO) to the Chief of Army Staff twice, while serving in the capacity of Military Secretary, responsible for HR management in the Army and as Director General Military Operations (DGMO) heading the Military Operation in the Army. During a career spanning four decades, General Bhatt has been the recipient of several awards from the Government of India. These include the Param Vishisht Seva Medal (2021) and other prestigious recognitions such as UttamYudhSeva Medal (2019), Ati Vishist Seva Medal (2015), Sena Medal (2014), Vishist Seva Medal (2012) and Chief of Army Staff Commendation Card (2008) for his dedicated service to the nation.

The Death of UCaaS:
Why AI-Driven Communications Are Eating the Collaboration Market

The death of voice-centric UCaaS and the rise of AI-driven platforms

The enterprise communications industry is undergoing a transformation as profound as the shift from hardware PBXs to cloud telephony. For the past decade, Unified Communications as a Service (UCaaS) dominated the narrative—bundling voice, messaging, video, and collaboration into standardized platforms. But that model is quietly being replaced by something fundamentally different: AI-driven, hyper-personalized communications fabrics built on programmable, API-first architectures.

The industry has spent the last decade building and scaling UCaaS. But businesses no longer want an all-in-one solution. They want customizable communication layers that fit their unique workflows, AI-driven real-time adaptability, and deep integrations that don't just connect calls but transform entire business processes. The future isn't "unified communications"—it's dynamic, AI-powered, and tailored to specific industries and business needs.

AI Co-Pilots for Work: The Agentic AI Revolution

The integration of generative AI into collaboration platforms represents the most significant shift in enterprise communications since the move to the cloud. What was once about connecting people is now about connecting intelligence.

From Chatbots to Autonomous Agents

The industry has moved decisively from generative AI to agentic AI—systems designed to autonomously analyze, strategize, and execute tasks with minimal human intervention. This is not incremental improvement; it is a fundamental redefinition of what collaboration platforms do.

Zoom has positioned itself as an "AI-first work platform," with its AI Companion now embedded across Zoom Docs, Meetings, Team Chat, Contact Center, Whiteboard, and Workflow Automation. The platform helps workers capture, interpret, prioritize, and act on information from multiple sources, live and asynchronously. Critically, Zoom's AI Companion has moved beyond basic tasks to agentic capabilities that can reason and automatically orchestrate events—scheduling meetings, tracking action items, and planning follow-ups.

Cisco's Connected Intelligence Framework

Cisco is making equally bold moves, positioning Webex as an "intelligent operating system for work". The company's Connected Intelligence framework is designed to carry context, knowledge, and automation across meetings, messaging, customer service, and workplace environments.

The centerpiece is a growing portfolio of AI agents intended to automate routine work. Cisco's Meeting Prep Agent automatically gathers information from previous meetings, documents, and message threads before a meeting begins. A Translator Agent provides speech-to-speech translation that preserves tone, emotion, and speaking style across 10 languages including English, Japanese, Mandarin, Spanish, German, and French.

Perhaps most significantly, Cisco has announced bi-directional integration between Webex and Microsoft Copilot, making Webex content searchable in Copilot and Microsoft content available in Webex. Cisco's AI Workforce Engagement Management platform was built specifically for environments that include both human and AI agents.

The Voice Resurgence Paradox

Ironically, AI may be breathing new life into voice—the very medium many declared dead. A 2024 survey found a quarter of people aged 18 to 34 never answer the phone, and nearly 70% of 18-34s prefer text to a phone call. Yet AI Receptionists—capable of taking voice calls, responding in real time, booking appointments, transferring calls, and providing business information—are eliminating the frustration of being on hold.

Zendesk reported that nearly 60% of customers feel long wait times are the most frustrating parts of a service experience, and 82% of respondents would talk to a chatbot if there were any waiting involved in talking to a human representative. AI Receptionists eliminate wait times altogether, potentially making voice a more attractive channel again.

After Enterprise Connect, where RingCentral demonstrated its new AI Receptionist product, its stock saw a notable uptick, breaking its downward trend and climbing away from the 52-week low. The message is clear: voice isn't dead—it's being reinvented by AI.

Network as a Service (NaaS): Networking's Subscription Model

Just as enterprise communications are moving from products to platforms, enterprise networking is undergoing a similar transformation. Network as a Service (NaaS) introduces a programmable, on-demand approach to wide-area connectivity that fundamentally changes how organizations consume networking.

The Promise of NaaS

Nile, a NaaSstartup founded by former Cisco executives John

Chambers and Pankaj Patel, has reportedly raised some $300 million to challenge the established networking order. Nile delivers a zero-trust network with no network operations required, offering a pure pay-as-you-use model—per user or square foot. "Think of it like Uber: you say where you want to go, and the service gets you there without you managing the vehicle". The company typically delivers a 40% to 60% reduction in total cost of ownership, including hardware, software, and lifecycle management.

The security model is particularly compelling. According to Gartner, 60-70% of cyberattacks originate at the networking surface area. Nile's architecture is designed to completely seal that surface area—no lateral movement is possible.

NaaS Adoption Trends

The market is responding. According to Wipro, a typical successful NaaS implementation results in a 35% reduction in network operational costs, a 60% improvement in deployment times, and a 45% decrease in security incidents. Gartner predicts that by 2028, 15% of enterprises will adopt on-premises NaaS, a significant increase from less than 2% in 2023. A Nokia and Pulse survey found that 47% of enterprise technology leaders plan to adopt NaaS, with 10% having already implemented such services.

The adoption pattern is instructive. Very large enterprises like JP Morgan Chase or Citi are not changing their core infrastructure overnight, but they are embracing NaaS at the edge—branch offices, retail locations, and remote sites where traditional IT support doesn't scale. This pragmatic approach allows enterprises to realize the benefits of NaaS without wholesale replacement of legacy infrastructure.

The Network Cloudification Opportunity

NaaS represents what industry observers call "network cloudification"—decoupling networking from legacy hardware and telco dependencies and delivering connectivity using cloud-native principles. Just as compute and storage moved into abstracted, scalable, usage-based services, networking is undergoing the same transformation.

For network engineers, this unlocks practical improvements: quickly provisioning temporary WAN paths for dev/test environments, automatically rerouting traffic for disaster recovery, integrating new sites in hours instead of weeks, and intelligently managing traffic for GenAI workloads with dynamic bandwidth allocation.

Private 5G: Vertical Case Studies and Real ROI

Private 5G networks are moving from proof-of-concept to production, with measurable ROI across multiple verticals. The technology promises dedicated resources that guarantee

performance, security, and ultra-low latency—capabilities that shared public networks simply cannot deliver.

The Economics: Build vs. Buy

The strategic decision to build a private 5G network or buy a 5G slice from a telecom operator is a critical techno-economic dilemma for industries requiring advanced automation. A comparative analysis of seaports, automotive factories, and power plants reveals significant trade-offs.

RAN slicing can reduce CAPEX by 33-68% and OPEX by 12-51% compared to private 5G deployments, primarily through macro-site reuse and shared resources. However, this comes at the expense of customization and the highest isolation and security levels. Industries that benefit from the customization and control offered by owning infrastructure must weigh these advantages against the cost savings of slicing arrangements.

Mining: Safety and Output Growth

Mining operations are among the most aggressive adopters of private 5G. The technology enables greater automation of production and more precise adjustment of volumes and timing based on market conditions. It is also a sector with high physical risk, where automation can significantly improve safety. Even without full automation, there are already use cases for worker protection in underground environments that become even more relevant with 5G due to real-time requirements.

Industry analysts estimate that edge computing combined with private 5G in mining and energy can deliver 5-8% output growth, translating to tens of millions annually in additional revenue.

Ports and Logistics: Turnaround and Revenue

Ports represent another prime opportunity. A container seaport with moving cranes and vessels requires guaranteed Quality of Service, enhanced security, and network flexibility. Faster turnaround at large hubs can translate to $50-100 million per year in additional revenue.

Morocco is positioning itself as a regional hub for private 5G deployment. Inwi, a Moroccan operator, announced a partnership with China Mobile International to deploy the country's first private industrial 5G network.

Manufacturing: Fewer Breakdowns, Faster Lines

Manufacturing leads all sectors with 272 private 5G deployments globally, according to GSA data. For automotive factories requiring sub-millisecond latency and massive device connectivity, private 5G enables Automated Guided Vehicle (AGV) systems and real-time production line monitoring. The result: 5-10% more output capacity without building new plants.

The value proposition is clear when considering the cost of failure. One disconnected automated guided vehicle due to lost signal can disrupt an entire production line. Private 5G's dedicated resources ensure critical operations remain connected.

Conclusion:

The Architecture of Modern Enterprise Communications

Three parallel transformations are reshaping enterprise communications:

From UCaaS to AI fabrics: The one-size-fits-all UCaaS platform is being replaced by dynamic, AI-powered communications layers that embed intelligence directly into business workflows. Voice isn't dead—it's being reinvented by AI, with agentic capabilities transforming everything from meeting preparation to customer service.

From hardware to subscriptions: NaaS is bringing the subscription model to networking, delivering a 40-60% reduction in total cost of ownership while eliminating operational overhead. The question isn't whether enterprises will adopt NaaS, but how quickly—and where in their infrastructure.

From public to private: Private 5G is delivering measurable ROI across mining, ports, and manufacturing—not as a theoretical concept, but as a commercial reality with real revenue impact.

The common thread is the shift from connectivity-first thinking to value-first, problem-solving approaches. Traditional vendors focused on selling voice, UCaaS, and connectivity face a stark choice: evolve beyond platforms to become true communications enablers, or risk being relegated to mere bandwidth providers.

The future belongs to those who see the broader shift from static UCaaS to dynamic, AI-powered communications fabrics. The enterprises that embrace this shift will dominate. Those that cling to the old model risk being disrupted.

From Connectivity to Trust:
How eSIM is Powering the Next Generation of Connected Vehicles

By Rahul Tandon, Senior VP, Connectivity Services, IDEMIA Secure Transactions

A decade ago, keeping a vehicle connected across borders required careful planning, region-specific SIM cards, and manual network management. Today, millions of connected vehicles move seamlessly between countries, updating software, transmitting diagnostics and enabling digital services without drivers ever noticing.

That invisible experience represents one of the biggest shifts taking place in the automotive industry.

Connectivity is no longer just about getting a vehicle online. It has become the foundation of the software-defined vehicle, where features, services and customer experiences continue to evolve long after the car leaves the factory.

For years, connectivity was a one-time decision. Manufacturers selected a mobile operator before production, installed a physical SIM, and that decision remained fixed throughout the vehicle's lifetime.

eSIM has fundamentally changed that model

Instead of locking vehicles to a single network, manufacturers can securely provision and manage connectivity profiles remotely using GSMA standards. Vehicles can receive new operator profiles over the air, adapt to regional requirements, optimise connectivity costs and maintain uninterrupted service throughout their lifecycle without any physical intervention.

The implications extend well beyond connectivity itself.

For automotive manufacturers, eSIM simplifies global production by enabling a single hardware configuration to serve multiple markets. Rather than maintaining different SIM variants for different regions, vehicles can be manufactured once and securely provisioned with the appropriate connectivity profile when they reach their destination. This reduces supply chain complexity while providing greater operational flexibility.

It also enables the digital experiences that customers increasingly expect.

Connected vehicles rely on always-on connectivity to support telematics, predictive maintenance, emergency services, navigation, over-the-air software updates and in-car infotainment. As vehicles become increasingly software-driven, reliable connectivity is no longer simply an added feature it is essential infrastructure that allows manufacturers to continuously improve the ownership experience throughout the vehicle's lifetime.

However, greater flexibility also introduces greater responsibility.

When connectivity decisions are made automatically rather than manually, every remotely provisioned credential must be authenticated, verified and securely managed. Trust therefore becomes as important as connectivity itself.

This is where secure connectivity management plays a critical role. Manufacturers require platforms capable of securely orchestrating eSIM lifecycle management at scale from profile provisioning and activation to remote updates and operator changes while maintaining the highest levels of security throughout the vehicle's operational life.

This challenge becomes even more significant when considering vehicle longevity. Cars manufactured today may remain operational for 10 to 15 years or more. Their security architecture must therefore be designed to evolve alongside emerging cyber threats, ensuring that connectivity remains resilient, trusted and compliant throughout the vehicle's lifecycle.

India is emerging as an important contributor to this evolution.

With rapid growth in automotive manufacturing, electric mobility and connected technologies, India is becoming both a major market and a global innovation hub for connected vehicles. As manufacturers increasingly build vehicles for global markets, scalable, secure and standards-based connectivity management will become a critical enabler of worldwide deployment.

Looking ahead, the conversation should move beyond simply connecting vehicles.

The real competitive advantage will come from creating trusted connectivity where every connection, every software update and every digital service is securely authenticated, intelligently managed and seamlessly delivered across the lifetime of the vehicle.

As vehicles continue their evolution into software-defined platforms, connectivity will increasingly become the invisible infrastructure that enables safer mobility, richer customer experiences and continuous innovation.

The future of connected mobility will not simply be defined by how vehicles connect.

It will be defined by how securely they stay connected.

Rahul Tandon

Rahul Tandon is Senior VP - Connectivity Services, India, IDEMIA Secure Transactions. Rahul has more than three decades of relevant experience in the telecom domain and is responsible for driving P&L delivering technology, products, and solutions to communications services providers (CSPs) and enterprise clients, and for defining and executing IDEMIA's telecom strategy across relevant audiences including OEMs, ODMs, Govt and telecom consumers. In 2017, he was selected to be IDEMIA's lead for its Mobile Operator Business.

How vRAN, RIC, and Multi-Vendor Deployments Are Rewriting the RAN Playbook

How Open RAN and vRAN are rewriting the rules of radio access networks

The radio access network—the foundational layer of mobile connectivity—is undergoing its most profound transformation since the transition from analog to digital. For decades, RAN infrastructure was the domain of a few vendors offering monolithic, proprietary systems where hardware and software were inextricably linked. Operators had little choice but to accept vendor lock-in, limited flexibility, and upgrade cycles measured in years rather than months.

That era is ending. Open RAN (O-RAN) and virtualized RAN (vRAN) are fundamentally reshaping how networks are built, operated, and optimized. By disaggregating hardware from software, opening interfaces between components, and embedding artificial intelligence directly into network control loops, these innovations promise to deliver the flexibility, cost efficiency, and intelligence that 5G and future 6G networks demand. This article examines three pillars of this transformation: the intelligent controllers driving real-time optimization, the energy-saving techniques addressing growing power concerns, and the real-world progress toward multi-vendor interoperability.

RAN Intelligent Controllers: The Brain of the Open RAN Revolution

At the heart of the O-RAN architecture lies the RAN Intelligent Controller (RIC)—a programmable platform that fundamentally changes how networks are managed. The RIC is split into two layers, each operating on different timescales to address distinct optimization challenges.

The Near-Real-Time RIC (Near-RT RIC) operates on control loops of 10 to 1000 milliseconds, hosting third-party applications called xApps that handle time-sensitive radio control tasks such as traffic steering, network slicing, and latency-sensitive optimization. These xApps can dynamically adapt resource allocation based on real-time network conditions, making decisions that would be impossible with traditional, manually configured systems.

The Non-Real-Time RIC (Non-RT RIC) operates on timescales of one second or more, orchestrating policies and analytics through rApps.

It provides high-level guidance and long-term optimization goals to the Near-RT RIC via the standardized A1 interface, enabling a hierarchical control structure that balances immediate responsiveness with strategic direction.

xApps in Action: AI-Driven Real-Time Optimization

The potential of xApps is being demonstrated through increasingly sophisticated implementations. Researchers have developed reinforcement learning-enabled xApps for near-real-time network slicing that orchestrate resources across diverse slice types—eMBB, URLLC, and mMTC—for multiple user equipments. These xApps dynamically adapt resource allocation and maintain quality of service under varying traffic demands, highlighting both the feasibility and challenges of end-to-end AI-driven optimization in disaggregated 5G frameworks.

The sophistication of xApp capabilities continues to advance. A framework called CORMO-RAN, implemented as anrApp, demonstrates lossless migration of xApps across compute nodes to save energy while ensuring service availability during migration. The orchestrator dynamically activates compute nodes based on xApp load and performs migration without causing RAN performance degradation. Experimental results on an O-RAN private 5G testbed showed energy savings of up to 64% compared to existing approaches.

The Agentic AI Frontier

Looking further ahead, researchers are exploring agentic AI frameworks that leverage large language models to enhance O-RAN resilience. An LLM-driven agentic AI framework has been proposed for real-time, adaptive resource allocation that dynamically interprets user equipment quality-of-service requirements and optimizes resource allocations across network slices.

Similarly, an edge agentic AI framework integrated into the RIC demonstrates proactive anomaly detection and autonomous network optimization. The framework achieved zero network outages under high-stress conditions, compared to 8.4% for traditional fixed-power networks and 3.3% for LLM agent-based approaches, while maintaining near-real-time responsiveness. These results establish that AI agents, when equipped with the right tools and contextual awareness, can be safely and effectively deployed in critical network infrastructure.

Energy Savings: Dynamic Cell Sleeping and Hardware Efficiency

As 5G networks expand and 6G approaches, energy consumption has become an urgent concern. Base stations account for 73% to 77% of total network energy usage, with a typical 5G macro base station consuming approximately 3.3 to 9 kW. The disaggregated nature of O-RAN creates both challenges and opportunities for energy optimization.

Dynamic Cell Sleeping with Machine Learning

The O-RAN architecture, with its RIC-based control, enables sophisticated energy-saving strategies that were previously impossible with monolithic RAN designs. Cell sleeping—switching base stations to low-power or sleep mode during off-peak periods—has been widely studied, but the O-RAN framework adds new dimensions of intelligence and automation.

Researchers have proposed deep reinforcement learning-based energy-saving solutions implemented as xApps that control the active/sleep mode of radio units from the Near-RT RIC. More scalable federated reinforcement learning solutions have been developed with an aggregator as anrApp in the Non-RT RIC and local agents as xApps. Results show that federated approaches achieve up to 43.75% faster convergence, more than 50% network energy saving, and 37.4% lower training energy versus centralized baselines, while maintaining quality of service and improving policy robustness.

Architectural Advantages of RIC-Based Energy Saving

The RIC-based energy-saving approach offers several architectural advantages that make it more effective than traditional methods :

Holistic visibility: The RIC has visibility across multiple radio units, enabling coordinated sleep and wake-up decisions that consider the entire network context.

Policy-driven flexibility: Operators can change sleep strategies simply by modifying policy specifications, without changing software.

Closed-loop operation: Cell sleep and wake-up decisions are based on key performance indicators, providing better network performance management and allowing the feature to be enabled continuously.

Future-proof design: Non-RT RIC and rApps are designed to work with both traditional RAN (via O1 adapters) and open/disaggregated RAN, enabling RAN automation evolution without "rip and replace."

Hardware Efficiency and Cloud Optimization

Beyond cell sleeping, energy savings can be achieved at the hardware level. Testing has demonstrated that controlling CPU power management features—dynamically turning off unused CPU components and adjusting frequency in response to traffic load—can save energy and reduce heat output while maintaining quality of service.

The impact of hardware accelerators such as ASICs, FPGAs, and GPUs in offloading CPU processing has also been demonstrated. In multi-vendor O-RAN 5G gNB implementations, peak throughput values close to theoretical limits were achieved while maintaining high virtual RAN workload efficiency, and peak-load MAC stress tests demonstrated scalability gains with hardware acceleration enabled.

Analytical power efficiency models for O-RAN reveal that radio units dominate the total network power budget, while distributed units benefit from sublinear growth due to statistical multiplexing, and centralized units scale the slowest owing to centralized control-plane sharing. The research identifies optimal ranges of RU-to-DU aggregation and accelerator utilization for maximizing energy efficiency in large-scale 5G and 6G O-RAN deployments.

Vendor Diversification: Multi-Vendor, Interoperable RAN Deployments

Perhaps the most tangible evidence of Open RAN's progress is the series of milestones achieved by operators moving from lab experiments to commercial deployment. AT&T has emerged as a bellwether for the industry's open RAN ambitions.

AT&T's Multi-Vendor Commercial Success

In August 2025, AT&T, Ericsson, and 1Finity (a Fujitsu company) successfully executed the first Open RAN call using third-party radios at AT&T Labs. This achievement demonstrated that disparate hardware—Ericsson's RAN Processor 6672 baseband and 1Finity radios—could interoperate through open interfaces.

Less than three months later, the same partners completed the first Open RAN call using third-party radios over AT&T's commercial network in Texas. This multivendor solution demonstrates the interoperability and flexibility Open RAN brings to our network, underscoring the critical role of collaboration in developing open and programmable networks that can transform the industry.

This commercial milestone is part of AT&T's ambitious $14 billion open RAN investment, announced in late 2023, with the goal of having 70% of wireless network traffic flowing through open platforms by the end of 2026. To date, AT&T has completed nearly 40% of its swap program from Nokia radios to Ericsson and has added mid-band spectrum to over 15,000 sites.

Third-Party rApp Deployment

AT&T also became the first operator globally to deploy a third-party RAN automation application (rApp) on a live production network. The rApp ran via Ericsson's Intelligent Automation Platform using the O-RAN Alliance's R1 interface, successfully optimizing live network performance. This deployment marks a shift from legacy single-vendor Self-Organizing Networks to open, multi-vendor programmable architectures, enabling operators to mix and match best-in-class automation tools.

Ecosystem Validation through PlugFests

The industry's progress toward interoperability is validated through initiatives like the O-RAN Global PlugFests, hosted by CableLabs and other organizations. These events test interoperability, efficiency, performance, and security of Open RAN components, enabling operators to build more cost-efficient, future-ready networks. The testing covers everything from CPU power management to hardware acceleration, building the confidence needed for commercial deployment.

The Pragmatic Path Forward

The open RAN movement is not without its challenges. O-RAN standards are still maturing, and O-RAN architecture-based RAN is not yet widely deployed. Some have questioned whether AT&T's reliance on Ericsson truly represents a multi-vendor strategy. However, industry analysts note that Ericsson's provision of RIC and service management software for non-Ericsson RAN equipment demonstrates that open RAN is accomplishing exactly what operators asked for—the ability to mix and match components from different vendors.

As Mobile Experts principal analyst Joe Madden observed, the shift toward open RAN represents the industry broadly moving in the right direction, however one defines it. The trajectory is now clear: networks are becoming more open, programmable, and intelligent.

Conclusion: The Open, Intelligent RAN

The RAN innovations driving today's networks represent a fundamental rethinking of how mobile infrastructure is built and operated. The RIC, with its xApps and rApps, is transforming network management from a reactive, manual process to a proactive, AI-driven capability. Energy-saving techniques, from dynamic cell sleeping to hardware acceleration, are addressing the growing power demands of 5G and future 6G networks. And multi-vendor deployments, like those demonstrated by AT&T, are proving that open, interoperable networks are not just theoretical constructs but commercial realities.

The journey from closed, monolithic RAN to open, virtualized, and intelligent infrastructure is still unfolding. Standards continue to mature, ecosystems continue to develop, and operators continue to gain confidence. But the direction is unmistakable: the future of the radio access network is open, programmable, and powered by artificial intelligence.

6G UNPLUGGED:

Terahertz Breakthroughs, RIS Tech, and the Standardization Countdown

The journey from research to standardization, and the breakthroughs that will define the next era of connectivity

As we approach the early 2030s commercialization target for 6G, the wireless industry is navigating a critical transition. The vision articulated by the International Mobile Telecommunications (IMT)-2030 framework is rapidly crystallizing into concrete technical specifications, with 3GPP having formally approved the timeline for the first 6G standard release.

Yet the path from research papers to live networks is neither straightforward nor revolutionary in the way some had anticipated. Instead, 6G is emerging as a pragmatic evolution of 5G's proven foundations, albeit one that introduces transformative capabilities in three key areas: the AI-native air interface, terahertz communications, and integrated sensing capabilities.

This article examines the current state of 6G standardization and the engineering breakthroughs required to turn its ambitious promises into reality.

The Road to Standardization: Pragmatism Meets Ambition

The 6G standardization journey is now firmly underway. Following the June 2026 3GPP RAN Plenary in Singapore, the Release 21 timeline has been clearly defined: the 6G Work Item is expected to commence in March 2027, with physical-layer functional freeze scheduled for September 2028, protocol design finalization in December 2028, and final ASN.1 freeze in March 2029 . From standardization completion to commercial deployment, the industry typically requires 12 to 18 months, positioning the first commercial 6G services for a 2029 to 2030 launch.

What's particularly noteworthy is the industry's deliberate conservatism at the physical layer. The foundational elements of 6G radio design—including OFDM-based waveforms, scalable numerology, LDPC and Polar coding frameworks—will largely inherit the proven foundations of 5G New Radio. This reflects a strong desire for ecosystem stability, cost-efficient migration paths, and backward compatibility.

"Our objective is to extend the system only where measurable gains can be demonstrated in spectral efficiency, energy efficiency, or user experience," explains Qualcomm's VP of Technical Standards, Juan Montojo. "6G is about coping with increased data demand while lowering the operator cost of ownership".

The research phase for 6G technology components began in mid-2025 and is expected to continue through early 2027. During this period, 3GPP's working groups are evaluating candidate technologies across multiple domains, including AI-native air interface design, integrated sensing and communication (ISAC) frame structures, and advanced channel coding extensions . The key challenge is balancing innovation with the practical constraints of implementation timelines, particularly for baseband hardware components like channel encoders and decoders that require lengthy design and validation cycles.

The AI-Native Air Interface: Intelligence Embedded in the Physical Layer

Perhaps the most consequential shift in 6G is the integration of artificial intelligence directly into the physical layer. In 5G, AI remained largely confined to management-plane tasks such as load balancing and power optimization. For 6G, this boundary is dissolving.

From Management to Signal Processing

The AI-native air interface represents a fundamental departure from traditional design principles. Instead of relying on standardized algorithms with fixed behaviors, 6G networks will employ trained neural networks for tasks such as channel estimation, modulation, and receiver processing. This approach promises site-specific optimization—the ability to adapt to diverse environmental conditions rather than applying one-size-fits-all solutions.

Research projects like the EU's CENTRIC initiative are advancing this vision by proposing a top-down, modular approach to wireless connectivity that places user communication needs and environmental constraints at the center of network stack design . "We use AI techniques to create and customize tailor-made waveforms, transceivers, signaling, protocols and RRM procedures to support these requirements," the project documentation explains.

The Pilotless Communication Breakthrough

One of the most compelling demonstrations comes from end-to-end learning approaches for the 6G air interface. Researchers have developed pilotless communication systems that eliminate traditional reference signals by jointly training neural receivers with custom trainable constellations. This approach embeds channel estimation mechanisms implicitly within the transmitted data, achieving competitive block error rates while delivering up to 8% higher goodput compared to 5G NR baselines—simply by reducing piloting overhead.

Scalable auto-encoder structures capable of supporting any M-ary modulation through a single AI/ML model have also been demonstrated, showing robustness against non-linear impairments expected in future high-frequency communications. These innovations require minimal modifications to existing infrastructure, providing a practical pathway toward intelligent, reconfigurable physical layers while maintaining backward compatibility.

The Validation Challenge

However, embedding AI in the physical layer introduces unique challenges. Unlike deterministic algorithms, neural network performance is tied to training history. A model trained on data from one urban environment might perform poorly when deployed in another. Engineers must ensure these models can generalize in the field, understand how models drift over time, and account for inherent biases.

Validation becomes essential. Techniques such as digital twins—exact 3D replicas of entire urban environments—are playing a vital role in demonstrating AI-native systems before deployment. Ray-tracing technologies model how signals bounce off buildings, trees, and moving traffic, allowing training within simulated environments before real-world rollout.

Terahertz Waves: The Engineering Challenge of the Spectrum Frontier

The terahertz band, spanning 0.1 to 10 THz, offers unprecedented bandwidth for multi-gigabit-per-second wireless links. Yet exploiting this spectrum requires overcoming formidable propagation challenges that have long kept it in the laboratory.

The Propagation Wall

Compared to lower frequencies, THz propagation suffers from severe atmospheric effects, high spreading losses, and molecular absorption losses that dramatically reduce transmission range and cell size. Reflection and scattering losses at THz frequencies severely attenuate non-line-of-sight components, resulting in extremely weak received power for reflected or scattered signals. The short wavelengths are also highly susceptible to blockage from environmental objects, with high penetration loss further limiting effective transmission range.

RIS: The Game-Changer

Reconfigurable Intelligent Surfaces (RIS) have emerged as the most promising solution to these challenges. These nearly passive metasurfaces with programmable phase shifts can steer and focus electromagnetic energy without adding active RF chains. At THz frequencies, where signals are highly susceptible to blockage, RIS can create controllable two-hop paths that restore line-of-sight-like links around obstacles.

Active RIS designs go further by amplifying reflected signals, effectively mitigating multiplicative fading effects. The engineering challenge extends to maintaining beam alignment—given the highly directional nature of THz signals—and managing discrete phase shifting, which is more practical for real-world deployment than continuous adjustments.

Recent research demonstrates that substantial SNR and energy efficiency gains are achievable by optimizing RIS configurations, even under severe fading and high path-loss conditions. Practical validation is advancing, with prototype designs achieving >90% reflection efficiency at 500 GHz and full 0-2π phase control. Beam-shaping experiments demonstrate single, dual, and multi-beam generation with high angular precision, establishing the practical viability of RIS for high-gain, low-power beamformingat THz.

Deployment Strategies

Given the short range and sensitivity to blockage, wide-area THz deployment is inefficient. Instead, researchers propose clustered heterogeneous network architectures where small base stations and users are spatially clustered around traffic hotspots. This approach allows THz communication to focus on providing high-throughput, low-latency services within localized areas such as stadiums, shopping centers, and schools. By strategically leveraging clustering and RIS, networks can mitigate the impact of short-range transmission while maximizing performance in dense, high-demand regions.

Holographic and Sensing Capabilities: Beyond Speed

Perhaps the most revolutionary aspect of 6G is the shift in its fundamental purpose. 6G is not merely about faster connectivity—it aims to create a network that perceives the physical world.

Integrated Sensing and Communication (ISAC)

ISAC has emerged as a foundational technology for 6G, motivated by the need for spectrum efficiency, hardware resource sharing, and the demands of bandwidth-hungry applications like haptic telemedicine and autonomous driving. The technology enables 6G systems to simultaneously transmit data and sense their surroundings, supporting applications from autonomous vehicles to smart city monitoring.

The integration is technically challenging. Sensing and communication objectives have contrasting requirements for waveforms and receiver design. At the transmitter side, novel ISAC beamforming and waveform techniques optimize the use of the same signal both for target detection and for serving multiple communication users. ISAC pilots that demonstrate good properties in estimating both communication channels and sensing targets have shown proven impact on probability of detection and channel capacity.

At the receiver, advancements in sensing estimation and imaging include hybrid radar fusion models that combine monostatic and bistatic sensing. Deep learning and parametric methods enable super-resolution joint estimation, while ISAC imaging modeling frameworks utilize ray-tracing-derived CSI to reconstruct 3D environments with high fidelity.

Microwave photonics is playing an enabling role, offering ultra-broad operating spectrum, rich parallel optical resources, and strong EMI immunity. Photonic frontend transceivers for 6G ISAC systems have demonstrated data rates up to 251 Gbit/s with sensing resolution from 0.25 to 750 cm.

Generative AI and Sensing

Generative AI is emerging as a complementary technology, addressing challenges such as data scarcity, network optimization, and personalized user experiences in ISAC systems. Traditional AI approaches to ISAC face limitations including inadequate data handling, poor generalization to different scenarios, and computational complexity. Generative models can alleviate these issues by synthesizing training data and adapting to diverse environmental conditions.

Security Implications

The convergence of sensing and communication introduces new security dimensions. ISAC systems that sense malicious eavesdroppers while ensuring communication secrecy create a dilemma that requires innovative solutions such as secure beamformingand artificial noise generation.

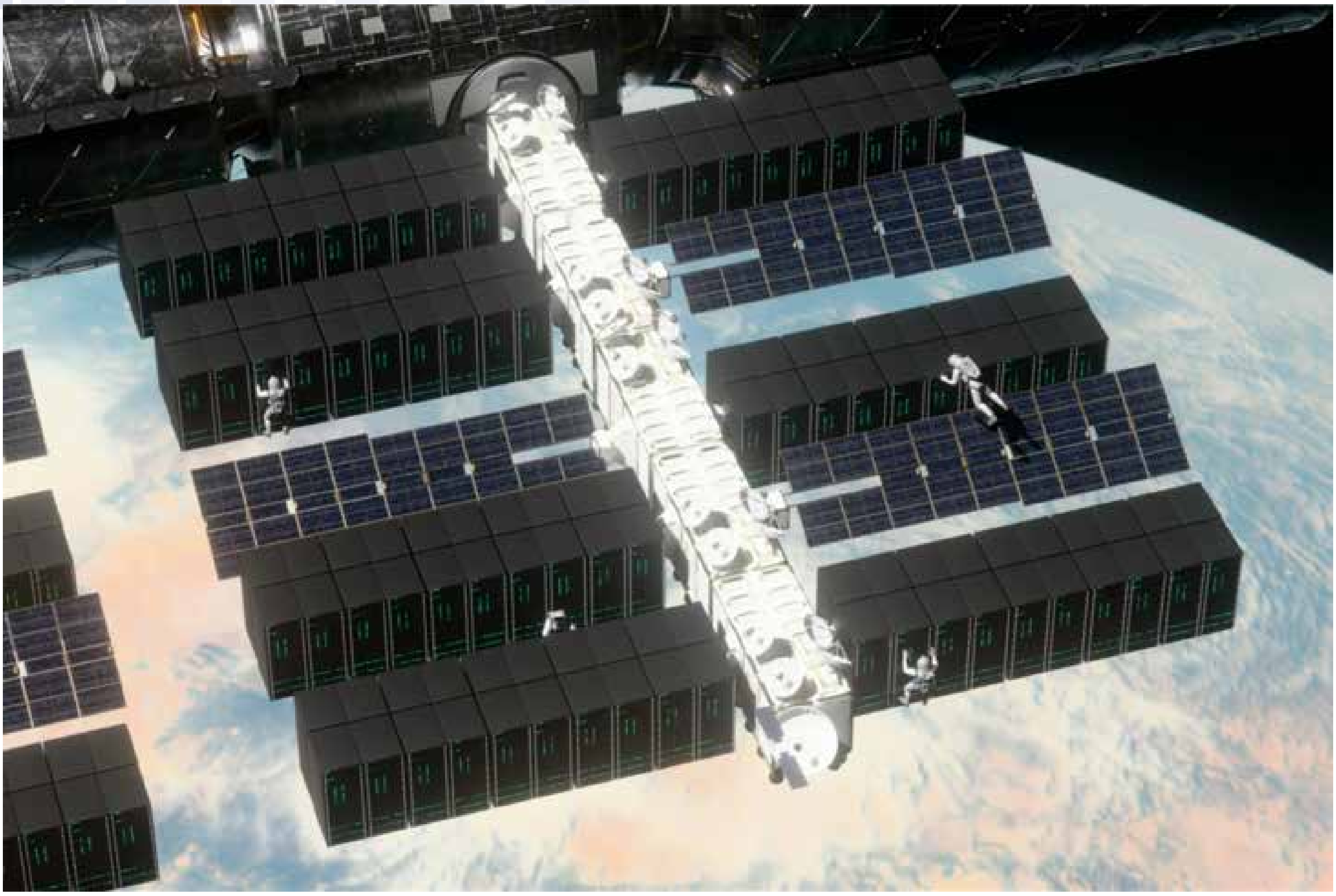
The Road Ahead

The transition from 5G to 6G is shaping up to be an evolutionary process that introduces revolutionary capabilities. The industry's pragmatic approach—retaining proven 5G foundations while introducing AI-native air interfaces, overcoming the propagation challenges of terahertz waves with RIS, and embedding sensing capabilities into the network fabric—reflects hard-won lessons from previous generational transitions.

As Cisco's leadership has emphasized, "95 percent of GPUs are going to massive hyperscalers to train models." The distinction for 6G is clear: it is not just the backbone of communications infrastructure but the nervous system of the AI era. The engineering breakthroughs required to realize this vision are substantial, but the roadmap is now clear and the timeline is set.

By the time the first 6G commercial services launch around 2030, the network will be capable not just of transmitting data at unprecedented speeds, but of sensing, perceiving, and intelligently adapting to its environment—fundamentally transforming what we expect from wireless connectivity. The standardization milestones achieved in 2026 mark the formal beginning of this journey from vision to reality.

AI Datacenters in Space: Hype Versus Reality



The concept of placing AI datacenters in orbit has rapidly shifted from science fiction to a tangible—if highly speculative—industry pursuit. As terrestrial datacenter expansion grapples with energy constraints, land shortages, and community opposition, tech pioneers and aerospace giants are looking upward.

But where does reality end and hype begin? This article examines current developments, the key players, and the genuine prospects for orbital compute infrastructure.

The Terrestrial Problem Driving the Space Solution

The push for space-based AI datacenters stems from an undeniable crisis on Earth. Ground-based datacenters face mounting constraints around energy, land, regulation, and scale, with the current model becoming increasingly difficult to sustain environmentally. In the past two years alone, community and environmental opposition in the US has stalled $180 billion worth of datacenter projects and delayed another $46 billion.

Proponents argue that space offers three fundamental advantages: near-continuous solar power without atmospheric loss or day-night cycles, passive radiative cooling into the vacuum of space, and elimination of freshwater consumption for cooling. Elon Musk has been characteristically bold, claiming at Davos that space-based datacenters could become the most cost-effective deployment location within two to three years.

Reality Check: What's Actually Operating Today

Axiom Space has emerged as the first to deploy genuinely operational orbital compute infrastructure. On January 11, 2026, the company launched two Orbital Data Center (ODC) nodes aboard Kepler Communications satellites. These nodes are not experimental prototypes but commercial service offerings, with workloads that can be submitted now.

Each node carries multi-GPU compute modules, onboard storage, and four optical terminals supporting 2.5 Gbps per link, compatible with

Space Development Agency Tranche 1 standards. This deployment followed a successful ISS prototype (AxDCU-1) in August 2025 that validated cloud computing and AI/ML inference using Red Hat Device Edge with MicroShift—essentially bringing Kubernetes container orchestration to orbit.

The Axiom deployment proves several things: commercial off-the-shelf (COTS) hardware can survive low-Earth orbit, containerized workloads can be deployed using familiar terrestrial tools, and AI inference and Earth observation data processing are viable at kilowatt-scale. What it does not prove is cost-competitiveness with terrestrial cloud or the ability to scale to large-model training workloads. Current optical intersatellite links deliver around 100 Gbps, while training large models requires approximately 7.2 Tbps of GPU interconnect bandwidth—one to two orders of magnitude short.

Key Players and Their Strategies

SpaceX: The Grand Vision

SpaceX has filed with the FCC for a constellation of up to one million satellites to support orbital AI computing. The company is fundamentally redefining Starlink from a broadband service to what it calls "connecting human infrastructure"—a global neural network linking AI systems, data, and intelligent terminals.

The company's proposed third-generation Starlink satellites are designed specifically for AI workloads. Each satellite would weigh 2,000 kg (compared to 575 kg for V2 models), operate in very low Earth orbit at approximately 350 km, and deliver a 10x improvement in downlink capacity (1 Tbps) and 22x improvement in uplink capacity (160-200 Gbps). Crucially, these satellites can only be launched via Starship, making the entire vision contingent on Starship achieving routine, high-frequency operation.

SpaceX has secured two massive compute lease agreements: a $9.2 billion monthly contract from Google for access to approximately 110,000 NVIDIA GPUs, and a $12.5 billion monthly deal from Anthropic. These agreements position SpaceX as an "AI Infrastructure-as-a-Service" provider, shifting from a rocket-and-satellite company to something far broader.

The company has also launched "Project TeraFab"—a joint venture with Tesla and Intel to build a 10-million-square-foot chip fabrication facility in Austin, with 80% of capacity dedicated to space computing and 20% to terrestrial applications.

Starcloud: The Dedicated Space Compute Startup

Formerly operating as Lumen Orbit, Starcloud has already sent an NVIDIA H100 GPU into space on a test satellite aboard a SpaceXFalcon 9 rocket. CEO Will Marshall claims this represented a 100x leap in compute power over any previous space-based computer. The company raised more than $10 million, operates within NVIDIA's Inception program, and claims a pathway to gigawatt-scale orbital capacity.

Starcloud's calculations suggest a 40MW datacenter in space would cost $8.2 million to deploy versus $167 million for a terrestrial equivalent over a decade. However, these figures rely on highly optimistic assumptions: $30/kg launch costs (currently ~$1,520/kg on Falcon Heavy) and solar array costs that exclude the protective glass required for space environments. The company envisions massive deployments requiring up to 4 square kilometers of solar arrays for a 5GW installation.

Blue Origin: Project Sunrise

Jeff Bezos' Blue Origin has submitted FCC plans for 51,600 datacenter satellites under Project Sunrise, with deployment of the TeraWave constellation slated to begin in Q4 2027. Bezos has described orbital datacenters as "very realistic" while questioning Musk's two-to-three-year timeline as "probably a little ambitious".

Google: Project Suncatcher

Google has partnered with Planet Labs on Project Suncatcher, planning to deploy 81 satellites equipped with Tensor Processing Units (TPUs) in sun-synchronous orbit at 650 km altitude. The project aims to create a space-based AI compute network using solar power and free-space optical communications, effectively bypassing terrestrial permitting bottlenecks.

Pixxel: India's Entry

Indian startupPixxel has partnered with AI companySarvam to develop the "Pathfinder" satellite, billed as India's first orbital datacenter, scheduled for launch as early as Q4 2026. The 200kg satellite will carry datacenter-class GPUs and Pixxel'shyperspectral imaging camera, enabling real-time AI analysis of Earth observation data directly in orbit. This mission focuses on sovereign AI infrastructure—running India-built models on India-built satellites to reduce dependence on foreign cloud platforms.

The Engineering Challenges

Cooling in a vacuum presents a fundamental paradox. While space is extremely cold, the near-vacuum environment makes convection and conduction almost ineffective for removing heat from AI chips. Existing solutions like those on the ISS require extremely heavy radiators, making launch costs prohibitive.

Radiation damage threatens hardware reliability. High-energy protons and particles can cause data corruption and hardware failure. While Google claims its Trillium chips performed well in proton beam tests, the survivability of commercial AI chips in the space radiation environment remains uncertain.

Space debris and collision risks escalate dramatically with constellations of hundreds of thousands or millions of satellites. Even with SpaceX's claim that atmospheric drag would deorbit failed hardware within approximately five years at low orbital altitudes, the collision risk from such dense constellations could render some orbital altitudes unusable.

Astronomical interference raises concerns from the scientific community. The Royal Astronomical Society estimates that if SpaceX's full plans materialize, approximately 10% of frames from the Very Large Telescope in Chile would contain satellite streaks.

The Economic Reality

The fundamental question remains: why do this in space when you could build on Earth cheaper? Harvard economist Mark Weinzierl, who studies space businesses, notes he has "never seen any [business models] that say right now [they're] cost-competitive". However, he suggests a convergence may occur: "the costs of doing it on Earth are going to go up over time and the costs of doing it in space are going to come down over time. And at some point those two curves are going to meet".

ABI Research estimates an orbital datacenter could cost more than 78 times a terrestrial equivalent, with Starcloud's third deployment projected as the first potentially cost-competitive option by 2028-2029. Even optimistic projections place meaningful deployment at five to ten years away at minimum—or up to 20 years for the kind of scale terrestrial datacenters have achieved today.

The Verdict: Hype vs. Reality

Hype: Orbital datacenters are already operational and economically viable.

Reality: Axiom Space has demonstrated operational compute nodes at kilowatt-scale, but these handle inference and Earth observation workloads—not training large AI models. They are not cost-competitive with terrestrial infrastructure.

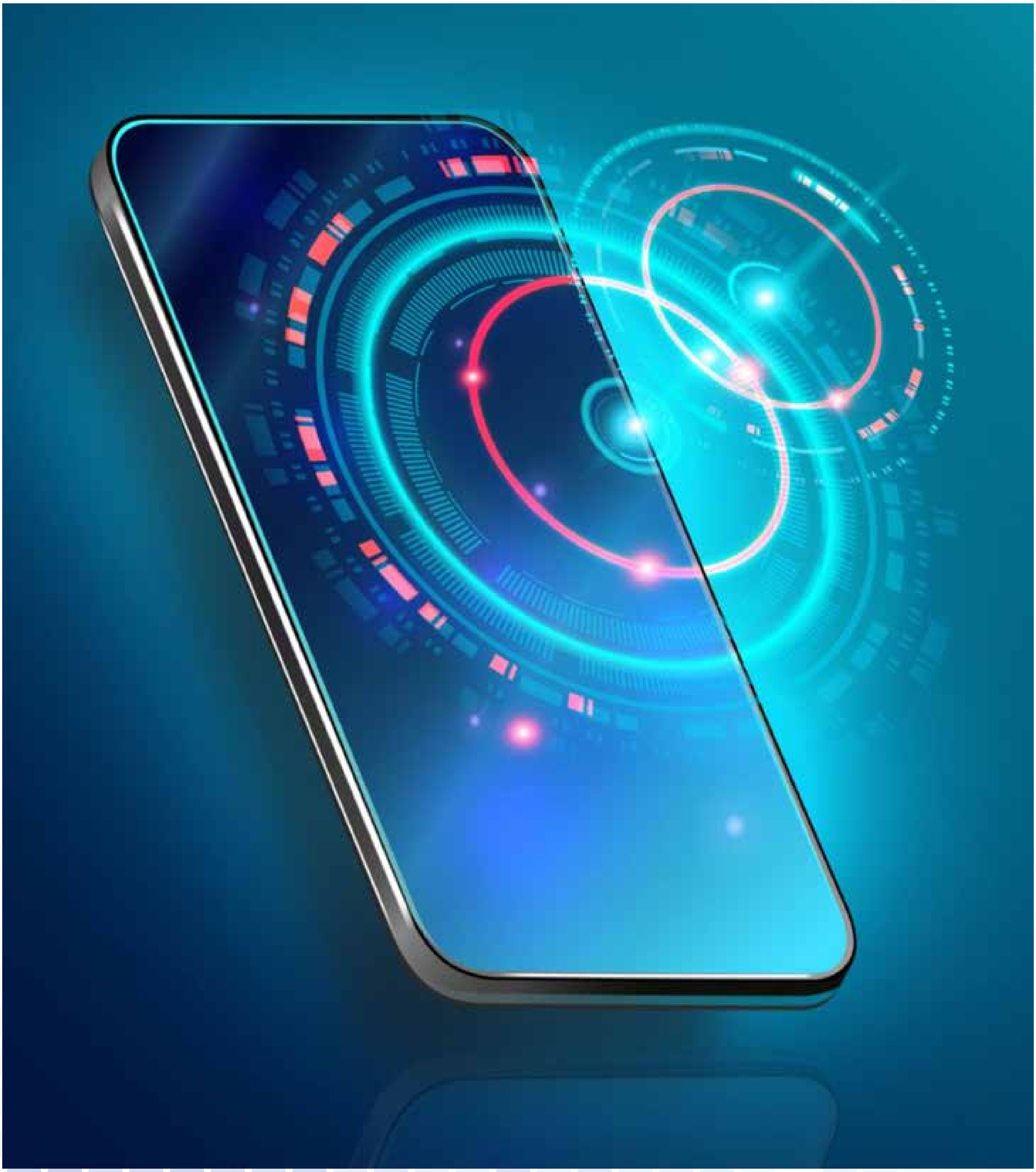
Hype: SpaceX will have a million-satellite orbital compute network within years.

Reality: The FCC application represents ambition, not operational reality. The timeline depends entirely on Starship achieving routine, high-frequency launches—which remains unproven. The application will face fierce regulatory opposition from competitors and agencies concerned about collision risks, orbital debris, and radio interference. FCC approval for the second-generation Starlink constellation took 18 months; this application faces far greater scrutiny.

What is real: The convergence of falling launch costs, advancing satellite manufacturing capabilities, and ever-increasing terrestrial datacenter constraints make orbital compute a plausible long-term development. The technology demonstrators are flying. The capital is flowing. But this is a decades-long play, not a two-year revolution.

For telecommunications operators, the implications are significant but distant. Space-based AI compute could eventually enable real-time processing of satellite imagery, edge AI for global IoT networks, and sovereign cloud infrastructure for data-sensitive applications. For now, watch the launch and testing milestones—particularly Starcloud's demonstrator missions, Axiom's expansion of its ODC network, and whether Starship achieves the launch cadence required to make these visions economically viable.

Future Smartphones | Exploring the Form Factor of the Post-Mobile Era



For over a decade and a half, the smartphone has reigned as the undisputed center of our digital lives. It is the first thing we check in the morning and the last we put down at night. Yet, a quiet but profound shift is underway.

The signs are everywhere: the market is saturated, innovation has plateaued, and the tech industry's most ambitious players are betting heavily on the next form factor. We are, perhaps, standing at the dawn of the post-mobile era—a world not devoid of smartphones, but one where they recede into the background, becoming the "hub" for a new generation of intelligent, ambient devices.

This transition is being driven by the confluence of several powerful forces: the maturation of Artificial Intelligence, the rollout of high-speed 5G and the promise of 6G, and the miniaturization of powerful computing components into ever-smaller form factors. The question is no longer if the smartphone will be supplanted, but how and when.

The Contenders: Glasses, Pins, and the Dawn of Ambient Computing

The most visible candidates to replace the smartphone are devices that don't look like phones at all. The world's leading tech companies are placing large bets on the vision of "ambient computing"—a model where technology is seamlessly integrated into our environment, anticipating our needs and acting proactively without requiring us to interact with a screen.

The market's momentum is undeniable. Counterpoint Research reported that in the first quarter of 2026, global shipments of smart glasses and smart head-mounted devices surged by 83% year-over-year. This growth is fueled by advanced AR glasses and simpler, screenless AI glasses like the Meta Ray-Ban. This category alone saw a massive 210% growth. The vision is for these glasses to become a constant companion, projecting navigation, messages, and contextual information directly into our field of view.

The hardware to support this future is rapidly evolving. Qualcomm, a key player in powering next-gen wearables, introduced its Snapdragon Wear Elite platform at MWC 2026. This chip is specifically designed for AI wearables and can run complex AI models with up to two billion parameters directly on the device, eliminating the need to constantly send data to the cloud. Meanwhile, Google is laying the software groundwork with its Android XR platform, which is designed to be the operating system for a new generation of extended reality devices.

The Engine Under the Hood: AI and Next-Gen Connectivity

These new devices are nothing without the intelligence and connectivity to power them. The "ambient" experience they promise relies heavily on "agentic" AI—artificial intelligence that can act as a personal assistant capable of reasoning and executing tasks across multiple devices without constant human intervention. A new system from Lenovo and Motorola, called "Qira," exemplifies this trend. It is a "Personal Ambient Intelligence" system that creates a persistent, context-aware AI layer across a user's entire ecosystem—from their smartphone to their laptop—ensuring your digital context is never lost as you move between devices.

This intelligence must be supported by a powerful network. Ericsson's Mobility Report highlights that AI-driven experiences like "see-what-I-see" sharing and multi-modal AI assistants will dramatically increase data traffic, especially in the uplink. 5G networks can handle the early versions of these services, but the true scaling of this data-intensive ecosystem will require the next generation of mobile technology: 6G. 6G is being designed specifically with AI-driven services and XR (Extended Reality) in mind, promising improved uplink performance, wider bandwidths, and more energy-efficient operation to allow for the widespread adoption of smart glasses and other personal AI devices.

What Does the Future Hold?

While the immediate future is likely to be an era of augmentation rather than replacement, the trajectory is clear. Smartphones will not vanish overnight; instead, they are expected to evolve into a powerful "hub" that manages connections to our glasses, watches, and other wearables.

We will reach for our phones less and less for mundane tasks like navigation or answering a text, as these will be handled by the device on our face or wrist. As a leading executive at OnePlus noted, the goal is to move beyond a platform for individual "walled garden" apps to one where AI connects different services to help users achieve tasks across them.

In the longer term, the potential applications are vast. Ericsson points to emerging industrial and enterprise use cases for AR, such as remote assistance where a technician receives step-by-step AR overlays on their equipment, or holographic collaboration where distributed teams interact with photorealistic 3D models. These experiences require not just great hardware, but a fundamental shift in network architecture to support low latency and high reliability.

Of course, this future is not without its challenges. Privacy and security remain paramount, as these systems are always on and constantly collecting data. The challenge for developers is to build trust by ensuring that consent is transparent and that as much data as possible is processed locally, on-device.

The "post-mobile" era is not the end of the smartphone, but the beginning of a new chapter in how we interact with the digital world. For now, the smartphone remains our "steadiest companion". But as the technology around us becomes smarter, more context-aware, and less intrusive, it may not be long before we find ourselves leaving that glossy slab of metal and glass in our pockets—where it belongs.

Beyond the Slab: Research Prototypes and the Future Smartphone

The smartphone may not be going away anytime soon, but the pace of innovation is certainly accelerating. While the headline-grabbing concepts of ambient computing and AI wearables suggest a post-mobile future, the reality is that the smartphone itself is undergoing a fascinating transformation. As highlighted at major events like Mobile World Congress (MWC) 2026, there is a clear "renewed experimentation in device hardware and interaction models," signaling that the industry is far from finished reinventing the slab.

Here's a look at some of the most intriguing research prototypes and what we can expect in the near future.

The Rise of "Embodied AI" in Phones

One of the most radical departures from the traditional smartphone is the concept of "embodied AI." This moves beyond a phone being a static, screen-based interface to one that can physically interact with its environment and user.

The standout example is the HONOR Robot Phone, a concept device unveiled at MWC 2026. This prototype integrates a robotic motion system directly into the chassis, utilizing a compact micro-motor and a four-degree-of-freedom gimbal. The device is designed to be aware of its surroundings through multimodal sensing, allowing it to track movement and maintain visual awareness. This enables truly novel capabilities:

All-angle AI Video Calling: The phone can physically track the user's position during a video call, ensuring perfect framing without the user having to hold the phone.

Expressive Interactions: The device can communicate through physical gestures like nodding or head shakes, creating a more lifelike and interactive layer to the user experience.

Next-Level Content Creation: A motorized arm on the back, topped with a 200MP camera, physically moves to stabilize videos and keep subjects perfectly framed, even when the phone is in motion.

While still a concept slated for testing in late 2026, the Robot Phone provides a powerful glimpse into a future where smartphones are not just smart but also responsive and physically adaptive.

The Evolution of Form Factors

Beyond adding robotic elements, the physical form of the smartphone is also evolving. Foldable phones are no longer the only cutting-edge design; we are now entering the era of "slidable" and "rollable" displays.

Samsung is a clear leader in this area, and journalists were recently given an exclusive look inside Samsung Display's headquarters to see its future concepts. Two designs stood out:

Flex Slidable: This device starts as a compact slab phone with a

4.7-inch screen. Through a sliding mechanism—possibly activated by pulling one end or pressing a button—a hidden display is pulled out, expanding the screen to a massive 7.2 inches. This allows for a device that is easy to carry but can transform into a tablet-sized screen for gaming or watching videos.

Flex Hybrid: This is an even more ambitious concept that combines two technologies for the first time: folding and sliding. The device can fold open and also features a sliding display on the right side to further expand its screen real estate. Samsung describes this as creating a product "that can satisfy both portability and entertainment needs simultaneously".

The question on everyone's mind is: when will these become a reality? Samsung's executive vice president noted that while the display technology may be ready, many engineering hurdles remain. He stated, "we should consider many aspects of the usage. When we have confidence that we can deliver meaningful innovation to users, then we are going to release the product". Despite this cautious approach, rumors suggest that a slidable Samsung phone could launch as early as 2027 or early 2028.

A New Type of Competitor: The SpaceX AI Device

The future smartphone market may also see new players. Reports indicate that SpaceX showed investors a prototype of a phone-like AI device during its IPO preparations. Described as being "thinner and more compact than the iPhone," the device is reportedly designed to run on a proprietary operating system and integrate AI technology from Musk's xAI. While Elon Musk has publicly denied these reports, stating they are "completely false," the rumors highlight the growing interest from non-traditional tech companies in

controlling the hardware through which users access their AI and connectivity services. The goal is reportedly to create a unified device that ties together communications, AI tools, and other services.

What to Expect in the Near Future

While the smartphone is evolving, a radical overnight replacement is unlikely. Instead, we can expect a period of augmentation. The smartphone will increasingly act as the "hub" for a network of wearables like smart glasses and earbuds, managing connections and handling complex computational tasks.

The immediate future of the smartphone itself appears to be driven by two key forces: intelligence and interactivity.

On-Device AI: Smartphones will have even more powerful AI capabilities processed directly on the device, enabling faster, more private, and more responsive AI assistants that can anticipate our needs.

Next-Level Imaging: The camera will remain a crucial battleground, with vendors leveraging AI for computational photography that rivals professional equipment and experimenting with physical, robotic camera systems like the HONOR prototype.

New Display Technologies: We will see the arrival of slidable and rollable phones, offering unprecedented flexibility in screen size within a compact form factor. Even the bezels around the screen are shrinking to nearly invisible levels.

In the near term, the "post-mobile" era is not an ending, but a transition. It is a journey where the smartphone is not rendered obsolete, but rather reborn as a more intelligent, adaptive, and central component of a much larger, more intuitive digital ecosystem.

Reimagining Future Smartphones

The future of the smartphone is being actively shaped by intense competition between the leading chipset manufacturers and device makers. While their strategies differ significantly, a common focus is emerging: making smartphones and connected devices more intelligent, context-aware, and adaptable. Here's what Qualcomm, MediaTek, and Samsung are doing on the smartphone research front.

Qualcomm: Pushing the Boundaries of Connectivity and On-Device AI

Qualcomm is focusing on two key areas to define the next generation of smartphones: advanced AI capabilities on the device itself and cutting-edge connectivity. A major announcement at MWC 2026 was the Snapdragon Wear Elite Platform. This new chip is a leap forward for wearables, which are increasingly seen as essential companions to smartphones.

Built on a 3nm process with a big.LITTLE architecture, it boasts a 5x improvement in single-core CPU performance and a 7x boost in GPU performance over its predecessor. Most importantly, it integrates a Hexagon NPU capable of running AI models with up to 2 billion parameters directly on the device. This enables powerful, low-latency experiences like real-time speech translation, an AI-powered health coach, and advanced gesture control—all without needing to constantly connect to the cloud.

For smartphones themselves, Qualcomm also unveiled new connectivity solutions at MWC 2026. The FastConnect 8800 and X105 5G Modem-RF systems are designed to power next-gen devices with AI-powered 5G Advanced and proximity-based AI experiences. This indicates a future where a smartphone can anticipate your needs based on location and context, seamlessly connecting to other devices and services. The first commercial devices using this new wearable platform are expected to launch in the coming months.

MediaTek: From Smartphones to a Broader AI Ecosystem

MediaTek showcased its vision for an "AI for Life" ecosystem at MWC 2026. While the company continues to power flagship smartphones with its Dimensity platform, its research extends far beyond that. A key area of focus is edge AI, where intelligence is processed on the device rather than in the cloud. For example, MediaTek

demonstrated its Dimensity 9500 chip being used in smart glasses to handle multimodal AI, processing both visual and voice inputs directly on the device.

MediaTek is also aggressively expanding into new form factors. A notable demonstration was a 5G NR NTN satellite video call from a vehicle, showcasing how satellite connectivity can keep cars and other devices "always online," a key step for ubiquitous connectivity. Furthermore, its new 3nm Dimensity Auto cockpit platform is designed to bring ray tracing, generative voice assistants, and gaming capabilities to vehicles, processing data locally for a more responsive and intelligent in-car experience. This strategy positions MediaTeknot just as a smartphone chipmaker, but as a key enabler for AI in a wide range of connected devices.

Samsung: Reimagining Form Factors and Expanding Galaxy AI

Samsung is taking a multi-pronged approach to the future of mobile, focusing on both new hardware and software intelligence. A major area of research is radically different smartphone form factors. Samsung Display demonstrated two "slidable" concepts: the Flex Slidable and the Flex Hybrid. The Flex Slidable starts as a compact 4.7-inch phone but expands to a 7.2-inch tablet-sized screen via a sliding mechanism. The Flex Hybrid is even more ambitious, combining a foldable hinge on one side with a sliding display on the other to offer unprecedented screen real estate and versatility.

While the display technology is ready, Samsung is cautiously approaching a commercial release, acknowledging significant engineering hurdles in mechanical reliability for slidable mechanisms. In the nearer term, Samsung is also refining its display technology with a 6.8-inch OLED panel featuring ultra-thin 0.4mm bezels, which could appear in Galaxy phones before the slidables.

On the software and hardware integration front, Samsung is heavily pushing its Galaxy AI ecosystem. At MWC 2026, the company showcased the Galaxy S26 series as the centerpiece of this strategy, highlighting features like an AI-powered Privacy Display, enhanced low-light photography, and a Photo Assist tool that lets users edit images with natural language commands. The company is also demonstrating new form factors like the Galaxy XR headset and the tri-fold Galaxy Z TriFold, indicating a commitment to exploring various screen-based interfaces.

In summary, these three technology giants are charting distinct but converging paths to the post-smartphone era. Qualcomm is building the powerful and intelligent chips that will power a new wave of AI wearables and phones. MediaTek is creating a broader AI ecosystem that spans from smart glasses to intelligent vehicles. Samsung is pushing the boundaries of hardware design and software integration to create a cohesive and intuitive "Galaxy AI" experience. All are betting on a future where our devices are more personal, more context-aware, and more seamlessly integrated into our lives.