International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

A Survey on AI and Cloud-Native Enablers for Automated Service Orchestration in 6G Networks

Johns Vincent
Wireless Networks And Applications Engineering
Amrita Vishwa Vidyapeetham
Kollam, Kerala, India

***Abstract*—6G networks aren't just another step forward-they promise to completely change how we connect and interact. Think holograms you can talk to in real time. Imagine the Tactile Internet letting you feel things remotely, or XR experiences that blend digital and physical worlds seamlessly. These futuristic applications bring a whole new level of complexity, on a scale the industry hasn't dealt with before. Handling this isn't something you can manage by hand anymore. We need networks that basically run themselves-zero-touch automation. This paper takes a close look at what makes that possible: Artificial Intelligence (AI) and Cloud-Native principles. First, we lay out the basics-what AI and Cloud-Native mean in this context. Then, we dive into how AI, especially Reinforcement Learning, is already stepping in to orchestrate services and manage network slicing on the fly. We also dig into the nuts and bolts of Cloud-Native architectures: the RAN Intelligent Controller (RIC), Multi-Access Edge Computing (MEC), and how non-terrestrial networks (NTNs) fit into the picture. These elements bring the speed and flexibility that automated systems need to actually work. Finally, we pull back to look at what still stands in the way: the lack of a unified management framework for hybrid networks (terrestrial and non-terrestrial) and the security issues that come with AI-driven control loops. Solving these is how we'll get to truly AI-native orchestration frameworks.**

***Index Terms*—6G, Network Orchestration, AI, Machine Learning, Cloud-Native, Network Slicing, Zero-Touch Management, Non-Terrestrial Networks (NTN).**

I. INTRODUCTION

Every new generation of wireless networks has stretched what's possible, but 6G is set to break the mold. Where 5G connects people and devices, 6G aims to dissolve the boundaries between physical, digital, and human worlds. Picture a network that's not just a data pipe, but a fabric for sensing, computing, and immersive experience-whether it's holographic calls, remote-touch surgery, or XR entertainment. Industries from healthcare to media will feel the impact. And this isn't just hype: most network operators (64%) say their main goal is to deliver these kinds of next-gen services. But moving from vision to reality isn't easy. 6G apps have brutal requirements. You need terabit-per-second speeds for holograms, sub-millisecond latency for tactile feedback, and "seven nines" reliability (99.99999%) if you're going to trust remote surgery to the network. Then add in the scale-up to 10 million devices per square kilometer. That's a management nightmare. Old-school, manual network management simply can't keep up. The real challenge isn't just bandwidth; it's handling the complexity, especially when current Quality of Service mechanisms start to buckle. For example, existing 5G RAN stacks often get bogged down by bufferbloat-greedy data flows clog things up, leading to huge queues and killing the low latency that time-sensitive apps need. Clearly, we need to move past manual intervention and shift to zero-touch automation. This article addresses that head-on, offering a unified survey of the two main forces driving this change: AI for smart decision-making, and Cloud-Native design for agile, adaptable architecture. While other surveys look at AI or Cloud-Native principles in isolation, we focus on their combined power as the engine for automated service orchestration as we move from 5G to 6G. Here's what you'll find in this survey:

- A big-picture look at network automation, tracking how management has evolved and spotlighting Cloud-Native and AI as the key technologies.
- A detailed taxonomy of how AI techniques tackle core orchestration challenges, from resource allocation to service activation.
- An analysis of major architectural frameworks like ETSI ZSM and O-RAN-that lay the groundwork for intelligent automation.
- A discussion of pressing open research challenges, outlining a clear path toward building a truly AI-native orchestration framework for 6G.

II. FOUNDATIONAL PILLARS

Before delving into the transformative influence of AI and cloud-native concepts on 6G, it's crucial to establish a solid understanding of the foundational principles that underpin these technologies. This section lays the groundwork by clarifying what "cloud-native" really entails, and by spotlighting the most impactful AI and machine learning approaches driving modern network orchestration and management.

A. The Evolution of Network Management

Network management has always been at the heart of telecommunications, encompassing a wide range of responsibilities: deploying resources, monitoring performance, configuring services, analyzing data, and maintaining control over the network's state. The overarching goals have remained constant-maximizing performance, reliability, and quality of service but the strategies and tools have evolved dramatically:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

- **Manual Management:** Initially, networks were operated through direct human intervention. Every configuration, troubleshooting step, and upgrade required skilled technicians to interact with hardware and software. This hands-on approach was not only time-consuming but also error-prone and difficult to scale as networks grew in complexity and size.
- **Scripted/Automated Management:** The next stage saw the advent of automation tools like SNMP and command-line interface (CLI) scripting. These allowed operators to automate repetitive and routine tasks, such as device monitoring and configuration updates. Automation made networks more efficient and reduced human error, but still required significant oversight and custom scripting.
- **Policy-Based Network Management (PBNM):** As networks expanded, it became clear that a rule-based approach was needed. Policy-based management introduced structured sets of rules that dictated network behavior in response to specific conditions. By offloading low-level management to policies, operators could focus on strategic objectives, while the network handled everyday operations automatically. This abstraction improved consistency and responsiveness.
- **Intent-Based Networking (IBN):** The real paradigm shift came with intent-based networking. Instead of specifying detailed step-by-step instructions, operators could simply define their desired outcomes such as a certain level of service quality or security posture-and the network would autonomously translate these intentions into concrete actions. IBN harnesses advanced analytics and automation to bridge the gap between high-level business goals and technical execution, freeing operators from micromanaging network processes.
- **Self-Organizing Networks (SON):** The introduction of SON, particularly with 3GPP Release 8, brought sophisticated automation to radio access networks. SON systems could autonomously handle tasks like cell planning, load balancing, interference mitigation, and fault recovery. While SON reduced operational overhead and improved adaptability, it often relied on static heuristics or predefined rules, necessitating periodic human intervention to fine-tune parameters and respond to unexpected situations.
- **Zero-Touch Network Management (ZTM):** The ultimate goal-and the focus of much current research and development-is Zero-Touch Management. ZTM envisions fully autonomous network and service management, where human operators oversee processes rather than intervene directly. By leveraging AI and machine learning, networks are now equipped to interpret their own status, predict future conditions, and make real-time decisions regarding planning, deployment, optimization, and healing. The ideal ZTM system can not only self-configure and self-optimize but also learn from experience, minimizing the need for manual oversight. Humans are only required to validate or approve critical actions, dramatically reducing operational costs and accelerating innovation. This survey will explore how AI and cloud-native architectures are converging to make ZTM a practical reality for 6G.

B. The Cloud-Native Paradigm

”Cloud-native” is far more than industry jargon-it represents a foundational shift in how modern networks are designed, deployed, and operated. Traditional, hardware-centric networks were rigid and slow to adapt, making them ill-suited for the dynamic demands of next-generation services. In contrast, cloud-native principles bring agility, elasticity, and resilience to the fore, all of which are essential for the hyper-connected, heterogeneous landscape envisioned for 6G. At the core of the cloud-native transformation is the Service-Based Architecture (SBA), which has become the backbone for both 5G and 6G core networks. SBA breaks down monolithic network functions into modular, reusable microservices. Each microservice encapsulates a specific capability-such as authentication, session management, or policy enforcement and communicates with other services through well-defined, standardized APIs. This modularity enables operators to rapidly introduce new features, scale individual functions as needed, and recover from failures with minimal disruption. Two pivotal technologies make the cloud-native paradigm possible:

1) **Microservices:** By decomposing network functions into granular, independently deployable units, microservices enable unparalleled flexibility. Each microservice can be developed, updated, and scaled in isolation, allowing for continuous improvement and innovation without risking the stability of the entire system. This approach also simplifies troubleshooting and maintenance, as problems can be isolated to specific services rather than affecting the whole network.
2) **Containers:** Containers further amplify efficiency by packaging applications and their dependencies into lightweight, portable units. Unlike virtual machines which require separate operating systems for each instance-containers share the underlying host OS, resulting in faster startup times and lower resource consumption. This efficiency is crucial for the dynamic, large-scale environments characteristic of 6G, where network functions need to be deployed, scaled, and migrated with minimal delay.

Effectively managing vast fleets of containers requires robust orchestration. Kubernetes has emerged as the industry standard for automating deployment, scaling, and management of containerized microservices. It provides powerful primitives for handling service discovery, load balancing, rolling updates, and fault tolerance. However, orchestration alone is not enough-microservices must also interact securely and reliably, often across distributed environments. A Service Mesh addresses this challenge by providing a dedicated infrastructure layer that manages inter-service communication. It enforces security policies, manages traffic routing, and collects telemetry, all transparently to the application itself.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

This enables operators to implement sophisticated networking features such as mutual TLS, circuit breaking, and traffic shaping-without altering application code. To support rapid innovation and maintain continuous service availability, cloud-native networks rely heavily on automated CI/CD (Continuous Integration/Continuous Deployment) pipelines. These pipelines automate the processes of building, testing, and deploying new services and updates, dramatically reducing the risk of human error and enabling networks to evolve at the pace of user demand and technological change. In summary, the fusion of AI-driven automation with cloud-native principles is setting the stage for a new era in network management-one defined by adaptability, efficiency, and self-sufficiency. As 6G approaches, these foundational pillars will be essential in meeting the unprecedented scale, complexity, and diversity of next-generation networks.

C. The AI/ML Toolkit for Network Automation

AI and Machine Learning (ML) drive network automation by letting systems learn from data, predict what's coming, and make smart, real-time decisions. When you look at network orchestration, the main ML techniques fall into three big groups:

- **Supervised Learning (SL):** Here, the model learns to map inputs to outputs using a labeled dataset-basically, it gets the right answers during training. The goal is to nail predictions or classifications on new data it's never seen before. In networking, people use SL for traffic forecasting, predicting when QoS or QoE will drop, classifying types of traffic, or spotting familiar security threats by their signatures. Popular algorithms? You'll run into Linear and Logistic Regression, Support Vector Machines (SVM), Decision Trees, Random Forests, and Neural Networks like Multi-Layer Perceptrons (MLPs).
- **Unsupervised Learning (UL):** This approach skips the labels. The model hunts for patterns or hidden structures in raw data, without anyone pointing out the ”right” answer. The aim is to uncover relationships or find oddities in the data. For networks, UL shines at anomaly detection-catching weird behavior or possible threats clustering users or devices by how they act, and shrinking down complex data with dimensionality reduction. Common tools here include K-Means Clustering, Principal Component Analysis (PCA), Autoencoders, and Gaussian Mixture Models (GMMs).
- **Reinforcement Learning (RL):** Think of this as learning by trial and error. An agent interacts with its environment, tries different actions, sees what happens, and gets a reward or penalty. The agent's job is to figure out a policy-a plan for what to do in every situation that racks up the most reward over time. RL is a natural fit for dynamic control and optimization, where the system keeps changing as it acts. You'll see RL in things like dynamic resource allocation for network slicing, adaptive traffic routing, power control, mobility management, and beamforming. Key algorithms include Q-Learning, Deep Q-Networks (DQN), Policy Gradient methods, and Actor-Critic algorithms. When deep neural networks join the mix, you get Deep Reinforcement Learning (DRL), which can handle huge, complicated state and action spaces.

These groups cover a huge range of algorithms-some of which you'll find summed up in the table below.

TABLE I
CLASSIFICATION OF COMMON ML ALGORITHMS FOR NETWORKING

Category	**Algorithm Example**	**Typical Network Use Case**
Supervised	Linear Regression	Predicting network load or traffic.
	Support Vector Machines (SVM)	Classifying traffic types or detecting anomalies.
Unsupervised	K-Means Clustering	Grouping users with similar behavior patterns.
	Principal Component Analysis (PCA)	Anomaly detection in performance data.
Reinforcement	Q-Learning/DQN Policy Gradient Methods	Dynamic resource allocation for network slices. Optimizing handover parameters in real-time.

III. ARCHITECTURAL FRAMEWORKS FOR AUTOMATION

To help the industry stay on the same page-and actually get things working together-several standards bodies have rolled out reference architectures for network automation.

A. ETSI Zero-Touch Service Management (ZSM)

The European Telecommunications Standards Institute (ETSI) ZSM group sketches out a blueprint for end-to-end, multi-domain automation. Their approach leans on a modular, Service-Based Architecture (SBA), which is built for easy scaling and future expansion. At the heart of this architecture are Management Domains, each shaped by either administrative control or technical boundaries. When a service stretches across more than one domain, an End-to-End (E2E) Service Management Domain steps in to handle cross-domain coordination. Each domain runs on closed-loop automation: Data Collection services keep an eye on the network, Analytics services sift through the numbers and spot trends, and Intelligence services decide when and how to act on managed resources. The Integration Fabric ties these domains together, making sure communication flows smoothly both inside and between them. To better visualize this, the ETSI ZSM reference architecture is depicted in Figure 1, which illustrates the modular, service-based approach with its distinct management domains.

B. O-RAN Architecture

The O-RAN Alliance takes a different angle, pushing for an open, intelligent, and virtualized Radio Access Network (RAN). Central to their vision is the RAN Intelligent Controller (RIC), which is split into two distinct parts:

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

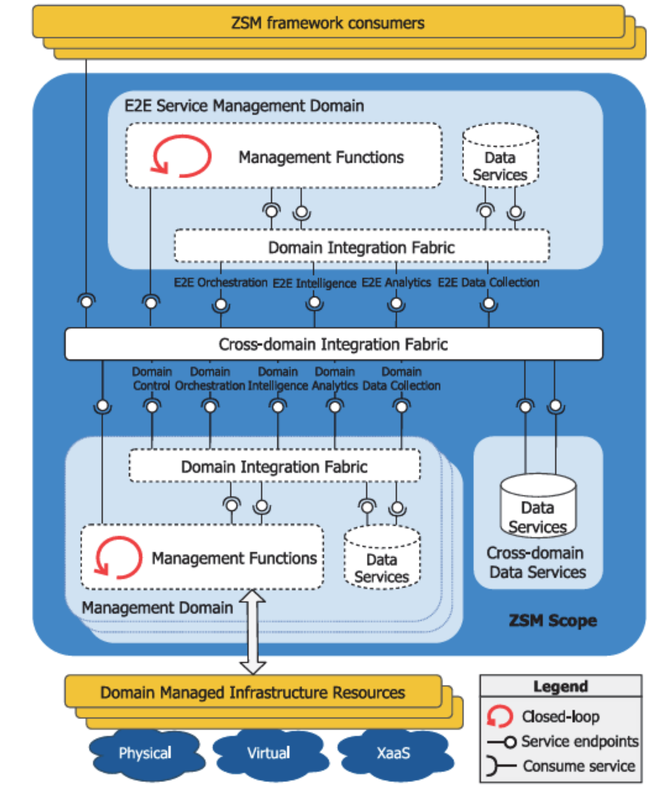

Fig. 1. The ETSI ZSM reference architecture, illustrating the end-to-end and domain-specific management loops (Source: Adapted from [3]).

- **Non-Real-Time RIC (Non-RT RIC):** Working at the SMO level, this platform manages policies and machine learning models for the RAN. It lets ”rApps” run control loops that don’t need to react instantly-anything slower than a second.
- **Near-Real-Time RIC (Near-RT RIC):** Deployed at the network’s edge, this platform hosts ”xApps” that interact with RAN nodes via the E2 interface, making decisions in near-real-time (from 10 milliseconds up to a second).

This layered setup means the system can handle everything from big-picture, long-term policy tweaks in the Non-RT RIC to rapid-fire, resource-level decisions in the Near-RT RIC. In practice, it creates several levels of nested control loops, each tuned for a specific timescale. Figure 2 provides a clear illustration of this layered architecture and its corresponding control loops.

IV. SURVEY OF AI-DRIVEN ORCHESTRATION AND AUTOMATION

Let’s dive deeper into how artificial intelligence and machine learning are reshaping the landscape of network orchestration and automation, especially as we move toward the era of 6G networks. The complexity of modern networks has grown dramatically they’re now highly dynamic, adapting constantly to changing conditions and diverse service requirements. Unlike traditional, static approaches where network

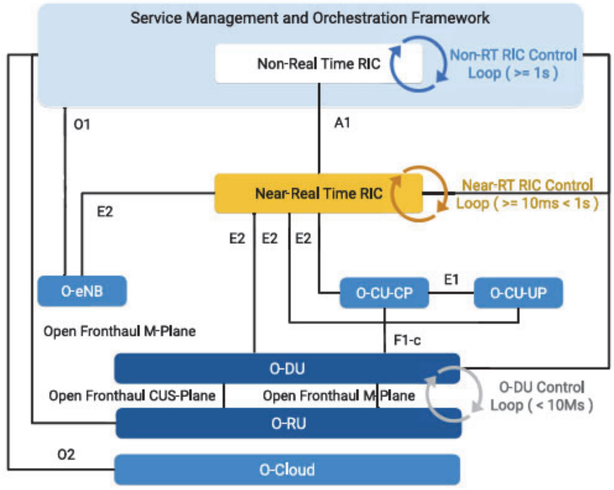

Fig. 2. The O-RAN architecture with its nested control loops, showing the interaction between the Non-RT RIC, Near-RT RIC, and the RAN elements (Source: Adapted from [3]).

resources were provisioned once and left untouched, today’s networks demand real-time intelligence and adaptability at every layer. Network slicing, which enables the creation of multiple virtual networks over a shared infrastructure, has intensified these challenges by introducing more granular and often conflicting resource needs. In this environment, static configurations quickly become obsolete. Instead, smarter, context-aware automation powered by AI is emerging as the only viable solution. AI-driven systems can make rapid, data-driven decisions throughout the entire lifecycle of a network slice: from initial deployment, through ongoing performance optimization as user demand fluctuates, to eventual decommissioning when the slice is no longer needed. The ultimate goal is to create autonomous networks that can intelligently allocate and reallocate resources such as compute power, storage capacity, and radio spectrum-on their own, even when multiple services with varying requirements are vying for those same resources. Researchers are at the forefront of this transformation, experimenting with a wide array of machine learning techniques to tackle the orchestration and automation hurdles unique to 6G and beyond. Their work is thoroughly documented in recent surveys, including those by Wajid and Rodrigo, which are highlighted in Table II. These surveys detail the latest breakthroughs, mapping out the specific network management problems addressed and the AI methodologies employed to solve them.

A. *AI for Dynamic Resource Allocation*

Dynamic resource allocation is fundamental to effective network slicing, as it directly impacts the network’s ability to meet a variety of quality of service (QoS) demands across diverse applications-from ultra-reliable low-latency communications required by autonomous vehicles, to massive connectivity for IoT devices, and high-throughput needs for multimedia

TABLE II
REPRESENTATIVE WORKS IN AI-DRIVEN NETWORK ORCHESTRATION

Key Innovation	Problem Addressed	AI Technique Used
AI-Native Slicing Framework: An integrated framework for space, air, and ground network segments.	Orchestration for 6G network slicing across different segments to meet diverse Quality of Service (QoS) demands.	A combination of deep reinforcement learning for resource allocation and recurrent neural networks for demand prediction.
Explainable MLOps (SliceOps): A framework to make the AI's decisions in network management more transparent and trustworthy.	Optimizing network slicing in 6G while ensuring the AI's decisions are understandable and reliable.	Explanation-Guided Reinforcement Learning (XRL) used to enhance dual-slice radio resource management.
AI-as-a-Service (AIaaS) Integration: Enhancing the standard NWDAF by integrating it with an AIaaS platform for more flexible ML model provisioning.	Overcoming the limitations of traditional 5G NWDAF architectures to improve anomaly detection and resource optimization.	A dual-slice approach separating "training slices" for continuous model updates from "inference slices" for real-time analytics.
Core and RAN Analytics Fusion: A system combining core network data (NWDAF) with real-time RAN data (RIC) for more effective resource control.	Dynamic resource allocation in 6G by integrating data from both the core network and the Radio Access Network (RAN).	A dual-slice paradigm where training slices refine ML models, and inference slices enable real-time network management decisions.
Dynamic Training/Inference Balancing: A framework to dynamically orchestrate resources to balance model training and low-latency inference.	Efficiently managing network resources by balancing the demands of data-intensive model training with low-latency inference needs.	AI algorithms dynamically manage resource allocation across separate training and inference slices.

streaming. Among the various AI methods explored, reinforcement learning and particularly deep reinforcement learning (DRL) stands out for its ability to handle sequential decision-making and adapt to real-time environmental changes.

- ***RAN Resource Allocation**:* In the radio access network (RAN), DRL agents are increasingly being used to allocate Physical Resource Blocks (PRBs) to different network slices as traffic patterns evolve. These agents process continuous streams of network data and QoS feedback, dynamically adjusting allocations to maximize efficiency and fairness. Compared to traditional static or heuristic-based schedulers, DRL-based approaches have consistently demonstrated superior adaptability and performance, learning to anticipate and respond to changing demands. Methods such as Q-learning, Deep Q-Networks (DQN), and Actor-Critic algorithms each bring unique strengths, allowing for more nuanced and customized allocation strategies that can be tailored to the specific needs of individual slices.
- ***Core Network Resource Allocation**:* The core network presents its own set of challenges, particularly in the placement and scaling of Virtual Network Functions (VNFs) or Cloud-Native Network Functions (CNFs) that are essential for supporting network slices. Here, DRL can be leveraged to make complex placement decisions, ensuring that latency-sensitive services are hosted close to the edge while optimizing the utilization of centralized cloud resources. By continuously learning from network feedback, DRL agents can anticipate bottlenecks and proactively migrate or scale functions, maintaining service quality even as network loads fluctuate.
- ***Edge Computing Resource Allocation**:* At the edge, the decision-making process becomes even more granular. AI systems must decide not only how to distribute computational workloads among local devices and edge servers, but also when to offload tasks to achieve the best balance between delay, energy consumption, and overall network efficiency. DRL has proven particularly effective in this context, enabling edge nodes to collaboratively learn optimal offloading and resource-sharing strategies that adapt to both local and global network conditions. This leads to reduced latency for end users and more sustainable energy consumption across the network.
- ***Beam Hopping/Management (SatCom/NTN)**:* In the realm of satellite communications and non-terrestrial networks (NTN), the challenges of resource allocation are compounded by the need to manage steerable beams and rapidly changing coverage areas. DRL agents are now being employed to dynamically allocate satellite beams based on real-time assessments of user demand and traffic distribution on the ground. These intelligent agents can prioritize coverage for regions experiencing surges in usage, optimize overall network throughput, and minimize latency, all while adjusting to the unique constraints of satellite mobility and coverage patterns.

Overall, the integration of AI and machine learning into network orchestration and automation is ushering in a new era of self-optimizing, resilient, and highly efficient networks. As these technologies mature, we can expect networks that not only react to current conditions but also anticipate future demands, providing seamless, high-quality connectivity for a wide spectrum of services and applications. The ongoing research and development in this space are paving the way for fully autonomous networks that will form the backbone of tomorrow's digital society.

B. AI for QoS and QoE Prediction/Assurance

Looking ahead-predicting how the network and user experience will change-lets operators act early and dodge SLA violations. Supervised learning dominates in this space.

- ***Traffic Forecasting**:* Models like ARIMA, LSTMs, and other RNNs predict future traffic-whether by cell, slice, or user. These forecasts steer resource allocation and scaling before spikes or dips hit.

- ***QoS/QoE Prediction**:* Supervised learning models estimate future QoS metrics or QoE scores using current network stats and traffic patterns. This gives orchestrators a heads-up, so they can fix issues before they affect users.
- ***Channel State Information (CSI) Prediction**:* In the RAN, deep learning models-CNNs and LSTMs-predict where channel conditions are headed. That's key for tweaking beamforming, adapting links, and making smart handover calls, especially when users move fast.

C. AI for Network Security and Anomaly Detection

AI steps up to spot threats and odd behavior in today's sprawling, complex networks. Unsupervised learning shines here, especially for catching new or never-before-seen attacks.

- ***Intrusion Detection**:* Deep learning models-Autoencoders, Deep Belief Networks (DBNs)-learn what normal traffic looks like, so they can flag anything that smells like trouble.
- ***Anomaly Detection**:* Unsupervised algorithms find patterns that don't fit the norm, signaling performance hiccups, misconfigurations, or outright security breaches.
- ***Traffic Classification**:* Deep learning can sort encrypted traffic based on flow patterns, dodging the need for deep packet inspection. This matters for both security checks and QoS management.

D. AI for Automated Service Activation and Provisioning

AI acts as a vital bridge between abstract service objectives and the intricate realities of network deployment, fundamentally transforming how networks are activated and maintained. Rather than merely facilitating initial configuration, AI-driven systems provide continuous oversight and adaptive management, ensuring that services remain optimized and resilient throughout their entire lifecycle.

- ***Intent Translation**:* Advanced machine learning algorithms analyze diverse, high-level service intents-such as latency, throughput, and security requirements-and dynamically translate them into granular network slice configurations. This goes beyond static parameter mapping, as AI can learn from historical data to refine how intents are interpreted, adapting to evolving business needs or regulatory changes. Such models can also resolve conflicts between competing service requirements, balancing priorities in real-time to maximize overall network efficiency.
- ***VNF/CNF Placement**:* The optimal placement of Virtual Network Functions (VNFs) and Cloud-Native Functions (CNFs) within a distributed, often multi-cloud infrastructure is a complex problem. AI enhances this process by considering a multitude of factors including current resource availability, predicted traffic patterns, and hardware constraints. It employs optimization techniques, sometimes in combination with reinforcement learning, to dynamically allocate functions where they will deliver the best performance, minimize latency, and reduce energy consumption. These intelligent placement strategies are crucial for meeting the stringent service level agreements (SLAs) required by emerging 6G applications.
- ***Slice Lifecycle Management:*** AI automates the end-to-end management of network slices, leveraging real-time analytics to make informed decisions about when to scale resources, conduct software updates, or gracefully decommission unused slices. Predictive models forecast demand surges and potential bottlenecks, enabling proactive resource provisioning that prevents service degradation. By integrating feedback loops and continual monitoring, AI ensures that network slices remain aligned with changing application requirements and user expectations, all while optimizing operational costs.

E. Analysis of Trends

The current landscape reveals a decisive shift toward more sophisticated, data-driven automation throughout network operations. Deep reinforcement learning is being harnessed for complex, real-time decision-making scenarios-such as adaptive resource allocation and fault recovery-where agility and precision are paramount. Meanwhile, supervised learning, particularly with advanced neural architectures like LSTMs and RNNs, is driving innovation in traffic and demand forecasting, resource usage prediction, and user behavior modeling. These predictive insights inform and enhance the automation pipeline. Unsupervised learning methods are gaining traction in network security, where they excel at uncovering anomalous patterns indicative of intrusions or failures without requiring labeled data. Additionally, hybrid approaches are emerging as particularly powerful: by coupling the foresight of supervised prediction with the adaptability of reinforcement learning, networks can both anticipate and respond to dynamic conditions more effectively. As AI increasingly assumes a central role in network management, the industry is placing heightened emphasis on explainable AI (XAI). Ensuring that AI-driven decisions are transparent and understandable is now critical for fostering trust among network operators and stakeholders, as well as for meeting regulatory requirements. Simultaneously, securing AI models-preventing data poisoning, adversarial attacks, and model theft-has become an urgent research priority. A notable development is the concept of Net4AI, which reimagines the network itself as an enabler for AI. In this paradigm, network architectures are purpose-built to facilitate large-scale, distributed AI workloads, offering specialized support for high-throughput data transport, edge inference, and collaborative learning. This approach is poised to create a virtuous cycle, where AI not only manages the network but is also empowered by it.

V. SURVEY OF CLOUD-NATIVE ENABLERS FOR A DYNAMIC 6G

While AI and machine learning lay the groundwork for intelligent automation, their true potential is only realized within a network that is agile, programmable, and inherently distributed. The cloud-native paradigm provides this essential foundation, introducing a suite of architectural components

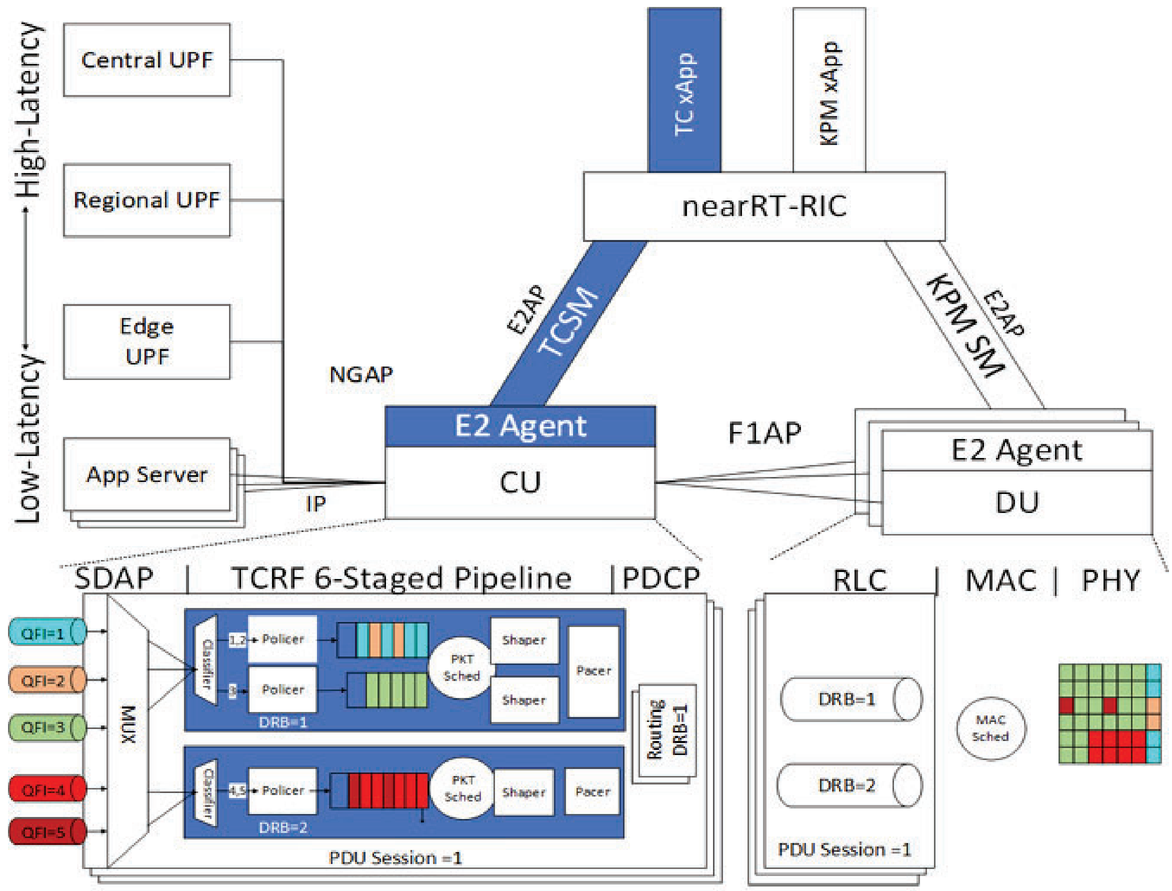

Fig. 3. An example of a disaggregated cloud-native architecture, showing the interplay between the nearRT-RIC, xApps, CU/DU split, and various UPF deployment models (Edge, Regional, Central) that enable dynamic service orchestration (Source: Adapted from [10]).

that unlock unprecedented flexibility and scalability. By abstracting network resources and exposing programmable interfaces, cloud-native design makes it possible for AI-driven control loops to operate in real time, orchestrating services and infrastructure with fine-grained precision. This section delves into the most transformative elements enabling this shift: from the granular programmability of the RAN Intelligent Controller to the distributed compute capabilities of Multi-Access Edge Computing (MEC), each plays a critical role in realizing the vision of a dynamic, responsive 6G ecosystem. A helpful illustration of how these components interact in a modern network is provided in Figure 3.

A. The RAN Intelligent Controller (RIC)

The RAN Intelligent Controller (RIC) represents a pivotal advancement in the evolution of radio access technologies. Integrated within the Open RAN architecture, the RIC fundamentally redefines how radio resources are managed and optimized. By modularizing the traditional, monolithic base station into disaggregated components, the RIC introduces standardized interfaces that empower trusted third-party applications-namely xApps and rApps to monitor, analyze, and control RAN behavior with unprecedented granularity and near real-time responsiveness. This open, programmable environment enables rapid innovation: AI-driven algorithms for spectrum management, load balancing, and interference mitigation can be deployed and updated dynamically, tailored to the unique demands of different deployment scenarios. The RIC not only enhances operational efficiency but also fosters a vibrant ecosystem of innovation, where diverse vendors and developers can contribute novel solutions. Ultimately, the RIC transforms the RAN from a rigid, vendor-locked domain into a flexible, intelligent platform capable of adapting to the fast-evolving requirements of 6G networks.

B. Multi-Access Edge Computing (MEC)

The advent of ultra-low latency applications such as tactile internet, immersive AR/VR, autonomous vehicles, and real-time industrial automation has underscored the need for computation to occur much closer to end-users. Multi-Access Edge Computing (MEC) addresses this challenge by distributing cloud computing resources to the network's edge, such as base stations, aggregation points, and enterprise premises. For orchestration platforms, MEC introduces a new dimension of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

complexity and opportunity. Decisions are no longer limited to how much resource to allocate, but now also encompass the optimal placement of workloads-evaluating whether tasks should execute in centralized data centers, regional edge clouds, or ultra-local edge nodes. This distributed computing fabric allows for rapid deployment and scaling of applications, directly in proximity to the devices and users that require them. By bringing compute, storage, and analytics capabilities to the edge, MEC not only reduces latency but also alleviates core network congestion, improves data privacy by limiting data exposure, and enables context-aware services that adapt to local conditions. For 6G networks, this flexibility is indispensable, empowering operators to deliver differentiated, low-latency services while supporting massive numbers of connected devices and diverse application requirements. The synergy between MEC and AI-driven orchestration further amplifies these benefits, enabling intelligent workload placement, proactive fault management, and seamless user experiences across the entire network continuum.

C. Non-Terrestrial Network (NTN) Integration

6G aims for global, everywhere-you-go connectivity, but terrestrial networks just can't stretch that far on their own. So, bringing in Non-Terrestrial Networks think LEO satellite constellations becomes essential [9]. By weaving together ground, air, and space, NTN integration builds a single, unified network. The orchestration framework faces a tough job here. It needs to keep services running smoothly, handle user movement, and juggle resources across domains that don't play by the same rules-latency and bandwidth swing wildly from one to the next. The management framework has to guarantee seamless connections, especially in remote places or disaster zones where ground networks vanish. All of this adds serious complexity to automating service activation.

D. Digital Twins

A Digital Twin is a virtual, living replica of a real-world system, process, or service. In networking, it acts like a high-precision simulation lab, echoing the true state and behavior of the live network.

- **Role in Automation:** Digital Twins give engineers a safe space to train, test, and validate AI-powered orchestration and control algorithms before rolling them out in the real world.
- **Implementation:** To pull off an accurate digital twin, you need real-time data feeds from the network, advanced modeling, and tight synchronization with physical systems.

E. AI Slice Architectures (Net4AI)

Net4AI pushes network architecture to evolve and truly support AI workloads.

- **Dual Slice Model:** Splitting network resources and QoS policies between AI Training Slices (built for raw throughput) and AI Inference Slices (tuned for ultra-low latency) lets the network serve both stages well.
- **Evolved UPF (e-UPF):** The UPF needs a boost-it has to recognize both slices and AI traffic, inspecting flows and steering them dynamically to the right MEC server, whether it's for data aggregation during training or for fast inference.
- **Integration with NWDAF/MEC:** The Network Data Analytics Function (NWDAF), or something like it, can work with the SMF and e-UPF to keep tabs on the lifecycle and resource needs of AI slices, making sure everything runs smoothly.

VI. OPEN RESEARCH CHALLENGES AND FUTURE DIRECTIONS

AI and cloud-native approaches have opened the door to smarter, more automated networks. Even so, our review of where things stand shows there's still a long way to go to reach a true zero-touch, fully orchestrated 6G network. The biggest hurdles? They cluster around automated service activation and provisioning.

A. Challenge: Lack of a Unified Framework for Service Activation Across Heterogeneous Domains

Everyone talks about a single, end-to-end management framework-look at the ETSI Zero-touch Service Management (ZSM) architecture [3]. In theory, it should handle everything. In reality, research is still split: RIC [10], cloud-native core [8], and NTNs [9] each get treated as their own separate problem. This gap shows up most clearly when you look at complicated service activation workflows. Imagine kicking off a smooth, low-latency XR session for a user speeding along on a train, jumping from 5G coverage to a satellite NTN link. That's a huge orchestration headache. You need one framework that can model performance, predict resource demand, and automate service activation and handoff across these completely different domains, all in real time. Right now, there's no proven orchestrator that can wrangle the wild swings in mobility and delay you find in NTNs, while also handling terrestrial MEC and RIC deployments, all under one roof. Until that exists, seamless service provisioning will stay out of reach.

B. Challenge: Security and Trust in AI-driven Activation Workflows

AI-powered automation has transformed network operations, but it also opens up a new vulnerability: the AI itself. Recent work points out that attackers don't just go after the network hardware or software anymore-they target the intelligence running the show. Data poisoning, evasion, inference attacks these aren't just buzzwords. They're real threats, and they exploit the fact that AI models can be manipulated. When the management AI receives poisoned data, it can make disastrous choices, all while thinking it's optimizing the network. Imagine an attacker feeding just the right malicious input, tricking the model into misconfiguring a network slice. Suddenly, there's a security gap, the promised quality of service gets trashed, or resources get drained,

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

causing denial-of-service for everyone else. The community knows about these problems, but we're still searching for practical, lightweight defenses that can actually protect the AI/ML pipeline during real-time network management and service activation. This is wide open territory for research.

C. Challenge: Explainable and Trustworthy AI (XAI)

Network operators aren't going to hand over the keys to an AI system unless they understand what it's doing. The problem? Most deep learning models are black boxes-you get an answer, but not the reasoning behind it. When something goes wrong, it's almost impossible to trace the cause. That's where explainable AI (XAI) comes in. XAI tries to bridge the gap by making AI decisions transparent and understandable to humans. But in practice, especially in the context of zero-touch management (ZTM), building models that networking experts can actually interpret-and developing metrics that measure just how explainable these models are is still a major challenge.

D. Challenge: Data Management and Federated Learning

Machine learning models live and die by their data. The richer and more representative the dataset, the better the model. But gathering huge amounts of network data and moving it to a central location? That's a privacy nightmare, especially when you're dealing with multiple vendors and operators. Federated Learning (FL) steps in as an alternative. With FL, the global model gets trained across many distributed devices, so the raw data never leaves its source. This protects privacy, but it's not a silver bullet. FL brings new technical headaches: communication overhead, the risk of attackers extracting sensitive information from gradients, and the challenge of working with wildly different data formats and qualities across devices.

E. Future Direction: Towards an AI-Native Orchestration Framework

Building an AI-Native Orchestration Framework These challenges make one thing clear: just slapping AI on top of existing cloud-native systems isn't enough for what 6G needs. We have to rethink the architecture itself. The orchestration framework of the future can't treat AI as an afterthought or a bolt-on module. It needs to be built around AI from the ground up, with intelligence, security, and explainability baked into the very core of orchestration and activation. This kind of framework unifies the management of diverse networks-terrestrial, non-terrestrial, you name it and raises the bar for what automated network services can do.

F. Towards an AI-Native Orchestration Framework

Tackling these problems means we can't just focus on AI for the network (AI4Net) or the network for AI (Net4AI) in isolation. The real challenge is designing an orchestration system where AI is a native part of the architecture-fully integrated with cloud-native principles, not just tacked on. The long-term goal is clear: develop an AI-native framework with unified data planes, intelligence embedded at every level, explainability and security designed in from day one, and seamless management of all kinds of resources, whether they're on the ground or in orbit. That's the vision for truly autonomous, intelligent, and trustworthy 6G networks.

VII. CONCLUSION

Getting 6G networks up and running-and keeping them running-will depend on how well we handle their complexity. Zero-touch automation powered by intelligent systems isn't just a nice-to-have; it's the only way to make these networks work at the scale and speed we expect. In this survey, we focused on the two pillars that make this shift possible: Artificial Intelligence, which brings learning and smart decision-making, and the Cloud-Native approach, which gives us the flexibility, scalability, and programmability modern networks demand. We started by breaking down the basics, showing how network management has moved from manual processes all the way to Zero-Touch Management (ZTM). Along the way, we mapped out the main cloud-native technologies and the AI/ML models shaping this evolution. Key standards like ETSI ZSM and O-RAN have set the stage for automation that actually works across different systems. When we looked at AI in orchestration, we saw Deep Reinforcement Learning dominating dynamic control tasks, Supervised Learning handling predictions, and Unsupervised Learning helping with security. Hybrid models and concepts like Net4AI-think dedicated AI slices are catching on, too. We also took a close look at 6G enablers like the RIC, MEC, NTN integration, and Digital Twins. Still, big gaps remain. Building a truly unified orchestration framework that works seamlessly across both terrestrial and non-terrestrial domains isn't just important it's essential. The AI components driving this automation need to be secure, trustworthy, and explainable. Managing the massive data flows that AI needs, while keeping privacy intact (possibly using Federated Learning), and making sure all this runs efficiently without burning through energy are major challenges we can't ignore. The next step is clear: we need an AI-native orchestration framework, built from scratch to serve as the secure, intelligent, unified core of future networks. Tackling these open research problems isn't just an academic exercise; it's the key to unlocking everything 6G promises-smarter, automated, and truly global connectivity.

REFERENCES

[1] K. Pradnya, S. Alam N. ”6G Wireless Networks: Vision, Requirements, Applications and Challenges”, *5th International Conference on Advances in Science and Technology (ICAST)*, 2022.
[2] J. S. Amin, Md. Jalil Piran, T. Amir, ”A Comprehensive Survey on 6G Networks: Applications, Core Services, Enabling Technologies, and Future Challenges”, *IEEE Transactions On Network And Service Management*, Vol. XX, No. XX, XX 2021.
[3] C. Estefania, S. Tejas. Adriana Fernandez, Md. S. Shuaib, Xavier Costa-Perez and Roberto Riggio ”Zero Touch Management: A Survey of Network Automation Solutions for 5G and 6G Networks”, *IEEE Communications Surveys & Tutorials*, Vol. 24, No. 4, Fourth Quarter 2022.
[4] K. Latif, Y. Ibrar, T. Nguyen, S.H Choong. ”Network Slicing: Recent Advances, Taxonomy, Requirements, and Open Research Challenges”, *IEEE Special Section On Green Internet Of Things*, Feb 2020.
[5] M. Qian, H. Fei and H. Qi. ”Deep Learning for Intelligent Wireless Networks: A Comprehensive Survey”, *IEEE communications surveys & tutorials*, vol. 20, no. 4, fourth quarter 2018.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 01 , January - 2026

[6] R. Wajid, R.B Joyeeta, F. Abraham, K. Diwakar, ”A Survey on Beyond 5G Network Slicing for Smart Cities Applications”, *IEEE communications surveys & tutorials*, vol. 27, no. 1, february 2025.
[7] M. Rodrigo, F.R.M. Larissa, O.S Flavio, ”Unleashing AI-Empowered Slices on Mobile Networks for Natively Cognitive Service Delivery”, *IEEE Access*, August 2025.
[8] B. Gabriel, ”Heavy Reading’s Cloud-Native SG Core Operator Survey”. White Paper on heavy reading. 2024.
[9] M.A Mahdi, S. Sourabh, C. Symeon, K. Oltjon, S. Hazem. C. Achiel, F.M.M Jesus, P. Sofie, H. Alireza, M. Arsham, L. Eva, O. Bjorn, ”Evolution of Non-Terrestrial Networks From 5G to 6G: A Survey”, *IEEE Communication Surveys & Tutorials*, Vol. 24, No. 4. Fourth Quarter 2022.
[10] I. Mikel and N. Navid, ”TC-RAN: A Programmable Traffic Control Service Model for 5G/6G SD-RAN”, *IEEE Journal On Selected Areas In Communications*, Vol. 42, No. 2, February 2024.