6G-AGENTS: Agentic AI for a Cognitive and Secure 6G Network-Compute Continuum

Behnam Ojaghi, Ricard Vilalta, and Raül Muñoz
Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Castelldefels, Spain
{bojaghi, ricard.vilalta, raul.munoz}@cttc.es
Mahdi Boloursaz Mashhadi, Mohammad Shojafar
5G/6G Innovation Centre (ICS), *University of Surrey*, Guildford GU2 7XH, United Kingdom
{m.boloursazmashhadi, m.shojafar}@surrey.ac.uk

***Abstract*—Next-generation 6G networks will demand radical advances in autonomy, scalability, security, and sustainability. This article introduces *6G-AGENTS*, a vision for embedding agentic Artificial Intelligence (AI) into the very fabric of 6G. We argue for a paradigm shift: from today's centralized, brittle automation toward decentralized Multi-Agent Systems (MAS) that operate as a native cognition layer, reasoning over intents, collaborating with Network Digital Twins (NDTs), and delivering self-optimizing, self-healing, and secure operations. We highlight opportunities, challenges, and an early use case in security, positioning agentic AI as a cornerstone of trustworthy, sovereign 6G infrastructures.**

I. Introduction

6G is expected to underpin immersive applications mainly those introduced by ITU-2030 [1] such as Immersive Communication (IC), Hyper Reliable and Low-Latency Communication (HRLLC), Massive Communication (MC), Ubiquitous Connectivity (UC), Artificial Intelligence and Communication (AIC), and Integrated Sensing and Communication (ISC), each aiming for different intent requirements. Meeting such requirements demands more than incremental automation; it requires a fundamental redesign of the network fabric to incorporate intelligence as a first-class citizen. As highlighted in standardization bodies such as ETSI [2], the complexity and scale of future networks require the development of highly autonomous, intelligent, and adaptive systems—capable of self-management, optimization, and dynamic evolution across both infrastructure and services. A key enabler of this transformation is the increasing viability of AI-native solutions, driven by advances in energy-efficient tensor computing. As the performance-per-watt of AI accelerators improves, running complex AI workloads becomes economically and environmentally feasible. At the same time, advancements in System-on-Chips (SoCs), CPUs, and GPUs are enabling sophisticated software-defined RANs that can fully exploit spectrum, cell sites, and energy budgets—paving the way for intelligent, software-driven network functions. Our project aims to address these challenges by developing Autonomous Cognitive Agents (ACAs) that will operate within a decentralized Multi-Agent System (MAS). These agents will form the foundation of a secure, scalable, and energy-efficient 6G architecture, capable of bridging the network-compute continuum with intelligence that is both context-aware and intent-driven. In particular, they will enable seamless service composition and knowledge graph handling across the network-compute continuum, allowing networks and applications to merge organically and dynamically to adapt to evolving conditions.

This necessitates a shift from statically configured networks to fully autonomous networks with *Self-X* capabilities—self-optimization, self-prediction, self-healing, self-configuration, and self-protection. The vision of 6G isn't just about faster speeds; it's about a cognitive network that understands the context of its operations, anticipates user needs, and autonomously adjusts to deliver optimal performance. This intelligence will be incorporated into every layer, from the radio access network to the core, enabling a seamless and personalized experience for a diverse range of applications and users. Furthermore, security and trust must be intrinsically linked to this intelligence, proactively identifying and mitigating threats in real-time, ensuring the resilience and reliability essential for the next generation of communication.

However, the transition to Agentic AI-driven networks presents significant challenges. Ensuring the trustworthiness and reliability of these autonomous agents is paramount. We need robust mechanisms for verifying their behavior, validating their decisions, and preventing malicious actors from exploiting vulnerabilities. This requires developing novel approaches to AI governance, incorporating principles of explainability, transparency, and accountability into the design of Agentic AI systems. Furthermore, the integration of diverse AI models and algorithms requires careful consideration to avoid bias and ensure fairness in resource allocation and service delivery. The cognitive plane must be carefully orchestrated to prevent chaos and ensure the emergence of coordinated, beneficial behavior.

Ultimately, the success of 6G largely relies on our ability to harness the power of Agentic AI responsibly and ethically. By embracing a human-on-the-loop governance model, we can ensure that these intelligent networks align with our values and priorities, creating a future where communication is seamless, secure, and truly transformative. This future demands a paradigm shift, one where networks are not just infrastructure, but intelligent partners, capable of anticipating our needs and empowering us to achieve more than we ever thought possible.

II. The 6G-AGENTS Vision

We envision a future where Agentic AI—autonomous cognitive agents operating in decentralized MAS—becomes the cognitive plane of 6G. This shift allows networks to move from human-in-the-loop control to human-on-the-loop governance, ensuring resilient, adaptive, and policy-aligned operations by introducing a new **cognition layer** as a core architectural element of the 6G system stack—alongside data, control, management, and orchestration layers. This cognitive plane will be natively responsible for real-time perception, learning, reasoning, and actuation across heterogeneous network domains, embedding knowledge graph processing and intent-driven intelligence as a native capability of the 6G architecture. This layer enables *Self-X* capabilities for network agents collaborating in a decentralized manner, as shown in Fig. 1. The cognitive layer is subdivided into two sub-layers: the intent sublayer and the agentic sublayer, enabling a full lifecycle of Intent-Based Networking (IBN) and agent-driven policy enforcement across services and resources. ~~The intent sublayer supports three functional stages—intent interface, intent fulfilment, and intent assurance—each mapped to dedicated IBN agents that enable autonomous interpretation, decomposition, execution, and SLA validation of service intents.~~ Intent interface acts as the entry point for service intents to interface and profile intent of services. Intent fulfillment translates and fulfills the specific intent requirements into actionable network tasks; Finally, intent assurance ensures that SLA/TLA of intent services are already satisfied using a closed-loop assurance approach. Intent sublayer enables advanced network features such as intent-driven slicing, dynamic scheduling, and Software Defined Network (SDN) control. innovates intent-driven slicing, scheduling, SDN control, etc. for example, service intents can be extracted, and key intent requirements can be mapped to the corresponding network slice. The agent sublayer serves as the "intelligent engine" of the network: deploying autonomous, collaborative agents that translate cognitive insights and intents into concrete, actionable operations, driving the Self-X vision of next-generation 6G networks. The addition of cognition layer simplifies the service exposure by implementing E2E Intent-Based Network (IBN) life-cycle automation, including intent service abstraction, intent translation and fulfilling and intent assurance, where the details will be discussed later.

6G-AGENTS isn't merely about automating tasks; it's about empowering the network with a form of distributed consciousness.

The decentralized nature of MAS is critical for the resilience and scalability of 6G. By distributing intelligence across the network, we avoid single points of failure and create a more robust system capable of withstanding unforeseen disruptions. Moreover, this decentralized architecture facilitates edge intelligence, pushing processing and decision-making closer to the users and applications, minimizing latency and maximizing responsiveness. These agents, trained on vast datasets and continuously learning from their experiences, will be capable of making decisions in real-time, adapting to changing conditions with unprecedented agility. We next detail the 6G-AGENTS architecture.

A. 6G-AGENTS Architecture

6G-AGENTS integrate advanced implementation of a MAS, Retrieval-Augmented Generation (RAG) orchestrated MAS, enhanced by the Agent-Based Interface (ABI), facilitating fully decentralized, intent-driven, and scalable management of 6G network operations. This architecture enhances self-healing, self-optimization, and adaptive decision-making by proposing a hierarchy of Agentic AI for 6G:

- **Agent-Based Interface (ABI)** a functional block for orchestrating communication among agents, tools and resources.
- **Coordinator Agents** that monitor global KPIs/KVIs and arbitrate among conflicting intents.
- **Domain-Specific Agents** that optimize tasks such as resource allocation, fault management, and security.
- **Predictive Agents** integrated with NDTs to validate "what-if" scenarios before live execution.

These agents operate in a perception–memory–planning–action loop, empowered by Retrieval-Augmented Generation (RAG) to ground decisions in real data and avoid AI hallucinations. The integration of Agentic AI within 6G networks promises a radical re-designing of network operations, moving beyond reactive management to proactive anticipation and intelligent adaptation. Consider a scenario where a sudden surge in demand occurs during a major sporting event. Instead of relying on predefined rules or manual intervention, Coordinator Agents, constantly monitoring network KPIs like latency and bandwidth utilization, detect the anomaly. These agents, informed by historical data and real-time sensor readings, can then intelligently arbitrate between competing demands, prioritizing critical services like emergency communications while dynamically allocating resources to support the increased traffic from the event.

Furthermore, Domain-Specific Agents would spring into action, optimizing specific aspects of the network to handle the increased load. Resource Allocation Agents, for instance, could dynamically adjust spectrum allocation, cell power levels, and backhaul capacity to ensure optimal performance. Fault Management Agents could proactively identify and mitigate potential bottlenecks, preventing service disruptions before they occur. And Security Agents, leveraging AI-powered threat detection, could identify and neutralize malicious activity, safeguarding the network against potential attacks exploiting the heightened activity.

Before implementing any significant network changes, Predictive Agents, integrated with NDTs, would simulate the impact of these changes on the network. This allows Mobile Network Operators (MNOs) to validate the effectiveness of the proposed solutions and identify any potential unintended consequences before they are deployed in the live network. Imagine the Resource Allocation Agent proposes a reallocation of spectrum to a specific cell site experiencing congestion. The Predictive Agent, using the NDT, could simulate this change and predict its impact on neighboring cells, identifying any potential degradation in service quality. This feedback loop allows the Resource Allocation Agent to refine its strategy,

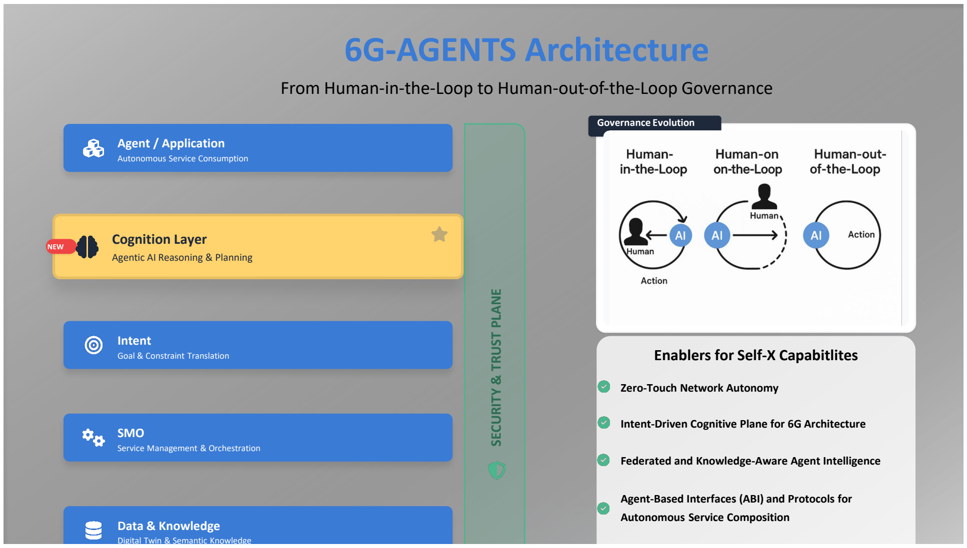

Fig. 1. The high-level illustration of the 6G-AGENTS.TBU

Figs/MAS.pdf

Fig. 2. 6G-AGENTS MAS Architecture.TBU

ensuring that the solution is both effective and minimally disruptive.

The RAG-powered *perception–memory–planning–action* loop is critical to ensuring the reliability and trustworthiness of these Agentic AI systems. By grounding their decisions in real-world data and constantly validating their actions against network performance, these agents can avoid the pitfalls of ”AI hallucinations” and ensure that their actions are aligned with the overall goals of the network. This continuous learning and adaptation is crucial for maintaining optimal network performance in the face of ever-changing conditions. Ultimately, the 6G-AGENTS vision seeks to create a self-optimizing, self-healing, and self-protecting network, known as *Self-X* capabilities, that can seamlessly adapt to the needs of its users, ushering in a new era of intelligent communication.

B. Opportunities

Embedding MAS into the network fabric opens multiple opportunities:

- **Simplicity**: translating high-level intents into automated actions reduces operational complexity.
- **Scalability**: decentralized collaboration among lightweight agents supports scaling to thousands of network functions.
- **Security**: proactive anomaly detection and intent-driven mitigation improve resilience.
- **Sustainability**: agents dynamically optimize resource usage, targeting 20–30% energy savings.

These opportunities position MAS as a key enabler for future networks, characterized by increased agility, efficiency, and robustness. Consider, for instance, the potential for:

- **Autonomous Network Slicing**: MAS could orchestrate and manage network slices dynamically based on real-time application demands and user requirements. Agents could negotiate resource allocation, optimize slice performance, and ensure isolation between slices, all without manual intervention.
- **Self-Healing Networks**: By constantly monitoring network performance and identifying potential issues, MAS can proactively trigger remediation actions. This includes automatic rerouting of traffic around failed links, scaling up resources in response to congestion, and implementing security patches to address vulnerabilities.
- **Personalized User Experiences**: MAS can personalize network services based on individual user preferences and behaviors. Agents can learn user profiles, predict their needs, and dynamically adjust network parameters to optimize their experience. This could involve optimizing bandwidth allocation for specific applications, prioritizing traffic based on user roles, or providing tailored security policies.

The realization of these opportunities requires a concerted effort in developing robust MAS frameworks, defining standardized agent communication protocols, and ensuring the security and trustworthiness of agent interactions. Furthermore, careful consideration must be given to the governance and control of MAS, ensuring that they operate within predefined boundaries and align with overall network policies. The journey towards fully autonomous, agent-driven networks is just beginning, but the potential benefits are substantial.

C. Challenges

However, several open challenges remain:

- **Trust and Explainability**: ensuring agents' decisions are transparent and verifiable under the EU AI Act.
- **Standardization**: converging on interoperable Agent-Based Interfaces (ABIs) and protocols for multi-vendor ecosystems.
- **Security Risks**: safeguarding agent-to-agent communications and preventing adversarial attacks.
- **Data Sovereignty**: enabling federated, privacy-preserving knowledge sharing across stakeholders.

These challenges necessitate a joint effort from researchers, developers, and policymakers. Addressing trust and explainability demands innovative approaches to agent design, incorporating mechanisms for self-explanation and auditability. This might involve the use of techniques like symbolic reasoning alongside neural networks, or the development of interpretable agent architectures. Consider the complexities of ensuring trust and explainability in a distributed agent system. How can we guarantee that an agent making a critical network decision is acting in accordance with predefined policies and not exhibiting unintended biases or vulnerabilities? Explainable AI (XAI) techniques can play a crucial role in providing insights into agent decision-making processes, enabling operators to understand the rationale behind their actions and verify their correctness. Furthermore, robust auditing and monitoring mechanisms are needed to track agent behavior and detect any deviations from expected norms.

Standardization requires a collaborative approach, fostering open-source initiatives and promoting the adoption of common ABIs. This will enable seamless integration of agents from different providers/vendors, unlocking the true potential of MASs. This could limit the scalability and flexibility of MAS deployments, hindering their ability to effectively manage complex network environments. Open-source initiatives and industry consortia can play a vital role in driving the development and adoption of standardized ABIs, fostering a collaborative ecosystem that promotes innovation and competition.

The inherent distributed nature of MAS also introduces unique security risks. Agent-to-Agent (A2A) communication channels could be vulnerable to eavesdropping, tampering, or impersonation attacks. Adversarial agents could be injected into the system to disrupt network operations or steal sensitive data. Robust authentication, authorization, and encryption mechanisms are essential to secure agent communications and prevent unauthorized access. Furthermore, anomaly detection techniques can be used to identify and isolate malicious agents, mitigating the impact of potential attacks.

Finally, data sovereignty presents a significant challenge, particularly in multi-stakeholder environments. Sharing knowledge and insights across different organizations is crucial for enabling effective collaboration and improving network performance. However, concerns about data privacy and security can hinder data-sharing initiatives. Federated learning and other privacy-preserving techniques can enable agents to learn from decentralized data sources without compromising the confidentiality of sensitive information.

D. Use Cases

The 6G-AGENTS vision extends beyond theoretical concepts and finds practical application in numerous real-world scenarios:

- **Smart City Traffic Management**: MAS can optimize traffic flow in real-time by dynamically adjusting traffic light timings, rerouting traffic around congested areas, and providing personalized navigation advice to drivers. Agents could collect data from various sources, including traffic sensors, cameras, and GPS devices, and collaborate to make informed decisions that improve overall traffic efficiency and reduce congestion.
- **Industrial Automation**: In industrial settings, MAS can enable autonomous control of robots and other machines, optimizing production processes and improving efficiency. Agents can monitor the performance of equipment, detect potential failures, and trigger maintenance actions automatically. They can also coordinate the movements of multiple robots to ensure smooth and efficient workflows.
- **Emergency Response**: During emergency situations, MAS can facilitate rapid and coordinated responses by dynamically allocating resources, optimizing communication channels, and providing situational awareness to first responders. Agents could collect data from various sources, including sensors, drones, and social media, and collaborate to create a comprehensive picture of the situation. They could then use this information to guide rescue efforts, allocate resources effectively, and provide timely updates to the public.

These use cases illustrate the transformative potential of MAS in enabling intelligent, autonomous, and resilient networks. As 6G deployments become more widespread, we can expect to see even more innovative applications of MAS emerge, revolutionizing the way we live and work. However, realizing this vision requires a collaborative effort from researchers, industry professionals, and policymakers to address the remaining challenges and ensure that MAS are developed and deployed in a responsible and ethical manner. The journey towards the Self-X network is complex, but the destination promises a future where networks seamlessly adapt to our needs, enhancing our lives and empowering us to achieve more.

III. Illustrative Use Case: Proactive Security Threat Mitigation

UoS Securing open, disaggregated 6G infrastructures is a critical challenge. In our vision, *Security Agents* continuously monitor telemetry, detect anomalies, and coordinate with *Self-Healing Agents* to apply countermeasures. Before enforcing actions, they consult an *NDT-Predictive Agent* to simulate impacts, ensuring continuity of service.

Preliminary experiments in controlled testbeds show that MAS-based mitigation can *halve mean-time-to-repair (MTTR)* compared to manual workflows. This highlights the potential of agentic AI to move from reactive recovery to proactive, preventive security operations.

IV. Future Directions

6G-AGENTS demonstrates that agentic AI is not only feasible but essential for future networks. Several key areas require particular attention in future research and development efforts: **Large-scale integration of agents across heterogeneous domains.** Large-scale integration demands robust agent coordination mechanisms and efficient resource allocation strategies. The complexities arising from managing thousands of agents across diverse operational contexts necessitate innovative solutions for conflict resolution and collaborative decision-making. Further, the agents must be capable of adapting to fluctuating network conditions and evolving user demands in real-time.

Federated learning and distributed knowledge graphs to ensure grounded, adaptive cognition. Federated learning and distributed knowledge graphs are crucial for equipping agents with the contextual awareness necessary for intelligent operation. By enabling agents to learn collaboratively without centralizing sensitive data, we can ensure both privacy and adaptability. Distributed knowledge graphs, representing the network's structure, capabilities, and real-time status, provide agents with a shared understanding of their environment, facilitating more effective decision-making and resource optimization. This approach also fosters resilience, as the loss of any single knowledge graph node does not cripple the entire system. **Alignment with standardization (ETSI ZSM, ENI, O-RAN) to foster adoption.** Alignment with existing standardization efforts is paramount for the widespread adoption of 6G-AGENTS. Integrating agentic capabilities within frameworks like ETSI ZSM, ENI, and O-RAN ensures compatibility and interoperability across different network deployments. This allows operators to seamlessly incorporate agentic functionalities into their existing infrastructure, accelerating the transition towards more intelligent and autonomous networks. By adhering to established standards, we can foster trust and confidence in agentic technologies, paving the way for their widespread deployment and utilization in future 6G networks.

AI-Native Agent Architectures: Moving beyond traditional agent models towards architectures that are intrinsically designed for AI integration is crucial. This involves exploring novel agent architectures that can seamlessly incorporate advanced AI techniques like deep learning, reinforcement learning, and generative models. Such architectures should be capable of continuous learning, adaptation, and autonomous decision-making in complex and dynamic network environments.

Explainable and Trustworthy AI for Agents: Ensuring that agent decisions are not only effective but also understandable and trustworthy is paramount. Research is needed to develop XAI techniques specifically tailored to MAS, enabling operators to understand the reasoning behind agent actions and verify their compliance with predefined policies. This includes exploring methods for generating human-readable explanations, visualizing agent decision-making processes, and providing guarantees about agent behavior.

Secure and Resilient Agent Communication: Protecting agent-to-agent communication channels from malicious attacks is essential for maintaining network integrity and security. Future research should focus on developing robust security protocols and mechanisms that can detect and mitigate various threats, including eavesdropping, tampering, and impersonation attacks. This includes exploring techniques like blockchain-based authentication, federated learning with differential privacy, and intrusion detection systems tailored to MAS environments.

Edge-Aware Agent Deployment: Optimizing the deployment and execution of agents across the network edge is crucial for minimizing latency and maximizing performance. This involves developing techniques for intelligently partitioning agent functionalities between the cloud and the edge, taking into account factors like network bandwidth, processing power, and data locality. Edge-aware agent deployment can enable real-time decision-making and control, enhancing the responsiveness and efficiency of 6G networks.

Integration with Digital Twins: Integrating MAS with digital twin technology can provide a powerful platform for simulating, analyzing, and optimizing network behavior. Digital twins can provide agents with a virtual representation of the network, enabling them to test new configurations, predict potential problems, and optimize performance in a safe and controlled environment. This integration can accelerate the development and deployment of innovative network services and applications.

These areas will unlock the full potential of 6G-AGENTS in transforming network management, paving the way for truly intelligent, self-optimizing networks capable of anticipating and adapting to evolving demands. Furthermore, the ethical dimensions of autonomous network management necessitate careful consideration. Bias mitigation strategies, eXplainable AI (XAI) techniques, and robust security protocols must be integrated into agent design to ensure fairness, transparency, and resilience. Exploring novel agent architectures, such as those leveraging neuromorphic computing or quantum-inspired algorithms, holds the potential to unlock unprecedented levels of performance and efficiency. Finally, a crucial step is the development of intuitive interfaces and management tools that allow human operators to effectively supervise and collaborate with these intelligent agents, fostering a symbiotic relationship between human expertise and AI automation in the 6G era.

V. Conclusion

Agentic AI heralds a disruptive shift for 6G: from static automation to a self-governing digital ecosystem. By embedding autonomous, collaborative agents into the network-compute continuum, we unlock a 6G infrastructure that is simultaneously more secure, sustainable, and sovereign. Standardized APIs and open-source platforms are critical for fostering innovation and enabling the widespread adoption of 6G-AGENTS. By establishing clear guidelines for agent behavior and data governance, we can ensure responsible and ethical development, promoting public trust and confidence in these transformative technologies. Moreover, comprehensive simulation environments and testbeds are essential for validating agent performance and identifying potential vulnerabilities before deployment in real-world networks.

The transition to agentic 6G networks also presents unique challenges in terms of workforce development. Training programs are needed to equip network engineers and operators with the skills necessary to manage and interact with intelligent agents. This includes developing expertise in areas such as AI, machine learning, distributed systems, and cybersecurity. By investing in education and training, we can ensure that the workforce is prepared to embrace the opportunities presented by agentic AI and contribute to the ongoing evolution of 6G networks.

Looking ahead, the integration of 6G-AGENTS with other emerging technologies, such as blockchain and the metaverse, holds tremendous promise. Blockchain can provide a secure and transparent platform for managing agent identities, access control, and data provenance. The metaverse can offer a rich and immersive environment for simulating and visualizing network behavior, enabling operators to gain a deeper understanding of agent interactions and optimize network performance. By exploring these synergies, we can unlock new possibilities for innovation and create a truly transformative digital ecosystem.

In conclusion, the journey towards agentic 6G networks is an ambitious but achievable one. By addressing the key challenges and embracing the opportunities outlined above, we can unlock the full potential of 6G-AGENTS and create a future where networks are not only more intelligent and autonomous but also more secure, sustainable, and beneficial to society as a whole. The future of networking is intelligent, collaborative, and undeniably, agentic.

REFERENCES

[1] ITU-2030, “ITU-R M.2160-0. Framework and overall objectives of the future development of IMT for 2030 and beyond,” 2023.

[2] ETSI, “ETSI GR ENI 051 V4.1.1, Experiential Networked Intelligence (ENI); Study on AI Agents based Next-generation Network Slicing,” 2025.