

Network Working Group
Internet-Draft
Intended status: Informational
Expires: 23 April 2026

E. Stephan, Ed.
Orange
R. Schott
Deutsche Telekom
D. Lopez
Telefonica
X. Duan
China Mobile
L. Morand
Huawei
20 October 2025

AI Agent protocols for 6G systems
draft-stephan-ai-agent-6g-02

Abstract

Communication between AI agents and between agent and tools is expected to be pivotal in 6G systems. The 3GPP TR 22.870 outlines various use cases and potential service requirements for AI agent communication within 6G systems. This document provides examples of use cases and service requirements contained in the 3GPP TR 22.870 and extrapolates possible requirements related to agent communication protocols.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 23 April 2026.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	3
2. AI Agent related use cases in the context of 6G	4
2.1. General	4
2.2. Network Optimization and Management	5
2.3. Immersive Communications	5
2.4. Hyper-Reliable Low-Latency Communications	6
2.5. Massive IoT Device Communications	6
2.6. Security and Privacy	6
2.7. Autonomous Systems	6
2.8. AI Agent Collaboration	7
3. Potential agent communications related requirements	7
3.1. General	7
3.2. Interoperability	7
3.2.1. Standardized Protocols	8
3.2.2. Multimodal Data Formats	8
3.2.3. Agent Identity Management	8
3.3. Discovery Mechanisms	8
3.4. Task Management	9
3.5. Context Awareness	9
3.5.1. Contextual Understanding	9
3.5.2. Adaptive Communication	9
3.6. Autonomy	9
3.6.1. Decision Making	9
3.6.2. Self-Management	9
3.7. Security	9
3.7.1. Authentication and Authorization	10
3.7.2. Data Protection	10
3.7.3. User Consent	10
3.8. Low Latency Communication	10
3.9. Reliability	10
3.9.1. Fault Tolerance	10
3.9.2. Load Balancing	10
3.9.3. Redundancy	11
3.10. Flexibility	11
3.10.1. Scalability	11
3.10.2. Adaptability	11
3.10.3. Extensibility	11

3.11. Energy Efficiency	11
3.11.1. Optimized Communication	11
3.11.2. Power Management	11
4. Conclusions	12
5. IANA Considerations	12
6. Security Considerations	12
7. References	12
7.1. Informative References	13
Authors' Addresses	13

1. Introduction

Since 1998, the 3rd Generation Partnership Project (3GPP) has been fundamental in the development of standards for various generations of mobile networks, including 3G, 4G (LTE), and 5G.

5G has revolutionized the way we connect, offering unprecedented throughput, low latency, and the capacity to handle a vast number of connected devices, thus driving innovation in the consumer market and various verticals such as healthcare, automotive, industrial automation, satellite and smart cities. Unlike traditional networks that rely on point-to-point interfaces, the 5G core network has been designed as a cloud-native service-based architecture with network functions communicating each other using RESTful APIs over HTTP/2. These network functions can be deployed and managed dynamically, leveraging cloud technologies such as virtualization, containerization, and microservices. This modularity allows for more agile, scalable, and efficient network operations.

Whereas these existing standards are still being enhanced to meet the growing demands of the telecommunications industry, the 3GPP has already undertaken a deep exploratory work on the use cases, service requirements and system architecture for 6G. This study phase will be then followed by a normative work to be completed by 2030 to meet ITU-R IMT 2030 timeline [M.2160].

6G aims to support societal advancements and to bring value to society in the 2030s and beyond in secure, resilient, environmentally and economically sustainable ways. In addition to new 6G services, other considerations are needed, e.g. CAPEX/OPEX reduction, improvement of overall 3GPP system performance, and migration from and interworking with 5G aspects.

A study on service requirements and use cases for 6G is documented in the 3GPP Technical Report (TR) 22.870 [TR22.870]. While at an early stage and the document being still a work in progress, the current content of the report already provides useful insights on the potential foundation pillars of the new 6G system. One of them being

the Artificial Intelligence (AI) and how 6G could leverage AI and machine learning to enhance mobile network capabilities, service offering and user experience.

2. AI Agent related use cases in the context of 6G

2.1. General

The recommendation ITU-R Recommendation M.2160-0 [M.2160] provides a comprehensive framework for the development of 6G technologies, focusing on the capabilities and objectives that these technologies should achieve around 2030 and beyond. In this document, the integration of AI in telecommunications is poised to be a cornerstone for the development of 6G systems. AI is considered as a foundational element, supporting both the network infrastructure and devices in delivering 3GPP services, often referred to as "AI for 6G system". Additionally, mobile network capabilities are aimed to be enhanced and optimized for supporting AI applications, termed "6G system for AI".

AI is not a novel concept within 3GPP, which has been actively engaged in standardizing AI and machine learning (ML) capabilities within 5G systems. These efforts span various domains, including management and orchestration, core networks, and next-generation radio access networks (NG-RAN). The objective is to boost system performance, efficiency, and the overall end-user experience. This ongoing work is crucial not only for current 5G advancements but also sets a solid foundation for future 6G technologies.

Beyond traditional AI/ML capabilities, 3GPP is exploring the integration of AI agents in 6G systems. According to 3GPP TR 22.870 [TR22.870], an AI agent is defined as follows:

AI Agent:

an automated intelligent entity that achieves a specific goal (autonomously or not) on behalf of another entity, by e.g. interacting with its environment, acquiring contextual information, reasoning, self-learning, decision-making, executing tasks (independently or in collaboration with other AI Agents).

AI agents are anticipated to enhance network efficiency by dynamically optimizing resources, predicting network conditions, and ensuring seamless communication between services. By incorporating large language models (LLMs), AI agents could interpret complex requests, convert them into actionable insights, and orchestrate advanced 3GPP services such as immersive communication, sensing, and computing services. It is also expected that these agents would be able to communicate, coordinate, and cooperate with other agents to tackle tasks that a single agent would struggle with.

In this context, communication between AI agents is expected to be pivotal in the 6G system, by enabling advanced network functionalities to enhance the existing capabilities of 5G networks, providing more efficient, reliable, and secure communication services.

The 3GPP TR 22.870 [TR22.870] outlines various use cases and potential service requirements for AI agent communication protocols within the 6G framework. Some of these use cases are provided in the following sections. They are only provided for information and are subject to change till the completion of the study.

2.2. Network Optimization and Management

- * **AI-Driven Network Performance Assurance:** AI agents in the 6G system should collaborate to ensure an efficient network configuration and resource allocation to provide on-demand 3GPP services at a given time and location area based on the authorized third party request by intent, e.g. such as for major sporting events or concerts in mega venues, etc. See clause 6.43 of 3GPP TR 22.870 [TR22.870].
- * **Predictive Maintenance:** AI agents deployed in the 6G systems should be able to share information to predict potential network failures and perform preventive maintenance, reducing downtime and enhancing network reliability. See clause 11.13 of 3GPP TR 22.870 [TR22.870].

2.3. Immersive Communications

- * **Traffic Management:** AI agents should be able to communicate with other agents and tools to analyze and predict traffic patterns, optimizing bandwidth allocation and ensuring high-quality service for users, especially in densely populated areas. See clause 6.52 of 3GPP TR 22.870 [TR22.870].

- * Quality of Experience (QoE) Optimization: AI agents should be able to collaborate to monitor and adjust network parameters in real-time, enhancing the user experience for high-bandwidth applications like video streaming and virtual reality. See clause 6.54 of 3GPP TR 22.870 [TR22.870].

2.4. Hyper-Reliable Low-Latency Communications

- * Real-Time Decision Making: AI agents should be able to facilitate real-time decision-making processes in critical applications such as autonomous driving, industrial automation, and remote surgery, where ultra-low latency and high reliability are crucial. See clause 6.3 of 3GPP TR 22.870 [TR22.870].
- * Fault Detection and Recovery: AI agents should be able to quickly detect and recover from faults, maintaining the high reliability required for critical applications. See clause 6.53 of 3GPP TR 22.870 [TR22.870].

2.5. Massive IoT Device Communications

- * Device Management: AI agents distributed in the network (terminal, edge, core, etc.) should be able to collaborate to manage a large number of IoT devices, optimizing their connectivity and power consumption to extend battery life and improve network efficiency. See clause 6.45 of 3GPP TR 22.870 [TR22.870].
- * Data Analytics: AI agents should be able to collaborate to share and analyze data from IoT devices to provide insights and support decision-making processes in various applications, such as smart cities, agriculture, and healthcare. See clause 6.45 of 3GPP TR 22.870 [TR22.870].

2.6. Security and Privacy

- * Anomaly Detection: AI agents should be able to exchange information about anomalous behavior detection and potential security threats in real-time, enhancing the overall security of the network. See clause 6.41 of 3GPP TR 22.870 [TR22.870].
- * Privacy Preservation: AI agents should be able to implement and share privacy-preserving techniques to protect user data, maintain sensitive data within individual networks, and ensure compliance with regulatory requirements.

2.7. Autonomous Systems

- * **Autonomous Vehicles:** Embedded AI agents should be able to communicate inside the same vehicle, between vehicles, between vehicles and the network for decision making based on environment perception, trajectory planning, and complex real-time control. See clause 6.3 of 3GPP TR 22.870 [TR22.870].
- * **Industrial Automation:** AI agents should be able to communicate to optimize and control industrial processes, improving efficiency and reducing operational costs. See clause 6.53 of 3GPP TR 22.870 [TR22.870].

2.8. AI Agent Collaboration

- * **AI agent communication groups:** Groups of User AI agents (AI assistant, drone, intelligent vehicle, home robot, etc.) could be dynamically created to allow the communication between AI agents to complete complex tasks requested by a user and/or between multiple user AI agent groups, owned by different users, across diverse locations ranging from local wireless networks to the wide area networks. See clause 6.7 of 3GPP TR 22.870 [TR22.870].
- * **Intelligent Communication Assistants:** Operators can advantageously provide Intelligent Communication Assistant services to their subscribers. Intelligent Communication Assistant could understand user intention by collecting multi-modal data of the user and execute the user instructions by invoking other AI assistants and services provided by the 3rd party service provider. See clause 6.11 of 3GPP TR 22.870 [TR22.870].

3. Potential agent communications related requirements

3.1. General

Agent communication in the context of AI and 6G networks involves several common requirements to ensure effective, efficient, and secure interactions between users and agents, between agents and between agents and tools. Here is a list of potential key requirements derived from the illustrative use cases provided in the previous sections. They are not yet formally approved by 3GPP and only provided for information/discussion.

3.2. Interoperability

3.2.1. Standardized Protocols

Agents should be able to use standardized communication protocols to ensure they can interact seamlessly across different platforms and systems. Agents should also support a standard protocol to interact with external data sources (e.g. user data repository) and AI-powered tools (e.g. location services) to complete their tasks.

3.2.2. Multimodal Data Formats

Agents should be able to support multimodal data formats (e.g., text, file, real-time audio stream, video streaming) to facilitate easy data exchange and interpretation.

3.2.3. Agent Identity Management

To allow the communication between users and agents as well as between agents and/or tools across platforms/domains, the system should support secure mechanisms for identification, verification and governance of agents/tools identities.

3.3. Discovery Mechanisms

Robust and efficient AI agent and tool discovery mechanisms to dynamically identify and locate AI agents/tools across different platforms and organizations. It should be possible to combine multiple discovery mechanisms depending on specific requirements and constraints of the application domain, e.g. broadcasting/multicasting-based mechanisms for intra-domain discovery and use of centralized directories for cross-domain discovery.

Currently Service-Based Architecture (SBA) defines discovery as the process by which Network Functions (NFs) locate and select other NFs that provide required services. This process is mediated by the Network Repository Function (NRF), which stores and exposes NF profiles containing service names, versions, and operational states.

The emergence of AI functions including autonomous agents, inference services, and retrieval-augmented generation (RAG) components, intend based exchanges, or training challenges this model. AI functions behave as both clients and servers within this ecosystem (but in 5G, almost every NF both exposes and consumes services too). These entities behave as application-layer servers that expose capabilities such as reasoning, summarization, optimization, or prediction, and as clients when they consume telemetry, policy, or contextual information from the network.

Supporting discovery for the various types of AI functions requires a broader analysis to identify solution-independent requirements.

3.4. Task Management

AI agent communication should provide a robust and efficient task management to enable seamless coordination and collaboration among agents. Task management includes task decomposition (complex tasks broken down into smaller sub-tasks that can be executed by individual agents or groups of agents), task assignment (based on agent capabilities, availability, workload), task scheduling, task coordination among agents and task monitoring/tracking.

3.5. Context Awareness

3.5.1. Contextual Understanding

Agents should be aware of the context in which they operate, including the state of other agents and the environment, to make informed decisions.

3.5.2. Adaptive Communication

Agents should be able to adapt the communication based on the context and current needs of the system.

3.6. Autonomy

3.6.1. Decision Making

Agents should be capable of making autonomous decisions based on data analysis and reinforcement learning but also making collaborative decisions based on dynamic interaction between agents.

3.6.2. Self-Management

Capabilities for self-management, including self-configuration, self-optimization, and self-healing based on information received from other agents.

3.7. Security

3.7.1. Authentication and Authorization

The agent communications should support authentication mechanisms for agent identity verification and authorization mechanisms to grant or deny access to specific resources based on the authenticated agent's permissions or roles. These mechanisms should be applicable for intra-domain and cross-domain scenarios.

3.7.2. Data Protection

The agent communications should support encryption mechanism to protect sensitive data from unauthorized access, ensuring privacy and confidentiality.

3.7.3. User Consent

The agent communication should provide mechanism to collect the user consent for the secure exchange of personal sensitive data with an AI agent (e.g. AI assistant) inside the network or 3rd-party application domain and between AI agents.

3.8. Low Latency Communication

The agent communication should enable minimal delay data transmission, especially critical for real-time applications like industrial automation, robotics, or real-time gaming, to reduce lag and improve service experience.

3.9. Reliability

3.9.1. Fault Tolerance

The agent communication should support mechanisms to detect, mitigate and recover from communication anomalies that would undermine collective decision making between agents. These anomalies include message storms, communication deadlocks, protocol violations, or content inconsistencies.

3.9.2. Load Balancing

Agent communication protocols should be able support load-balancing among agents to prevent any single agent from becoming a bottleneck and maintain continuous operation.

3.9.3. Redundancy

Redundancy mechanisms (e.g. active/passive redundancy, data redundancy, georedundancy) should be implemented to ensure high-reliability, resilience and service continuity even if one or several agents fail.

3.10. Flexibility

3.10.1. Scalability

Agent communication protocols should be designed to accommodate increasing network sizes and data volumes without significant performance degradation, even during peak load events.

3.10.2. Adaptability

Agent communication protocols should be able to adjust to multiple contexts (including heterogeneous agent capabilities, service environment, system requirements), multiple ways of transmitting information, including verbal, non-verbal, written, and visual forms, used individually or in combination, as well as the location of the AI agents (in the user terminal, in the network, in the 3rd party service environment).

3.10.3. Extensibility

For onward compatibility, agent communication protocols should be able to evolve and adapt to new technologies and service requirements without requiring major overhauls.

3.11. Energy Efficiency

3.11.1. Optimized Communication

Agent communication should support mechanisms to optimize communication between embedded AI agents to reduce energy consumption, especially important for battery-powered devices (e.g. IoT devices).

3.11.2. Power Management

Agent communication should support efficient power management strategies to extend the operational life of devices embedding AI agents.

4. Conclusions

AI agents are envisioned to enhance network efficiency by dynamically optimizing resources, predicting network conditions, and facilitating seamless communication between services. The incorporation of large language models (LLMs) will enable AI agents to understand complex requests and orchestrate advanced services, further enhancing the capabilities of 6G networks.

AI agent communication is expected to play a crucial role in enabling advanced network functionalities. The use cases outlined in 3GPP TR 22.870 [TR22.870] demonstrate the potential of AI agent communication to enhance the existing capabilities of 5G networks, providing more efficient, reliable, and secure communication services.

In summary, the integration of AI in 6G systems represents a significant advancement in telecommunications technology. The ongoing work by 3GPP in standardizing AI capabilities and exploring the potential of AI agents highlights the transformative impact that AI is expected to have on future network infrastructures and services.

If a multi-AI agent-based system is formally adopted by 3GPP in the scope of 6G, standard solutions will be required to support secure and reliable communication between agents and between agents and external tools. These solutions will be used inside the 3GPP system but also with 3rd-party platforms. It is then required to have solutions developed by standard organizations that would be widely adopted by the AI development community. It is foreseen that IETF could be the right place to develop and maintain such standard protocols. If such standardization work is eventually endorsed by IETF, a close coordination between IETF and 3GPP will be essential to ensure that any AI agent communication protocols specified by IETF will support specific functional and service requirements defined by 3GPP in the 6G context. And it is also expected that this work will be completed in a timely manner to cope with the challenging workplan defined by 3GPP for the development of a 6G system in the ITU-R IMT 2030 framework [M.2160].

5. IANA Considerations

This memo includes no request to IANA.

6. Security Considerations

This document should not affect the security of the Internet.

7. References

7.1. Informative References

- [TR22.870] 3GPP, "3GPP TR 22.870: Study on 6G Use Cases and Service Requirements; Stage 1 (Release 20)",
<https://www.3gpp.org/ftp/Specs/archive/22_series/22.870>.
- [M.2160] ITU-R, "Recommendation ITU-R M.2160-0: Framework and overall objectives of the future development of IMT for 2030 and beyond", <https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2160-0-202311-I!!PDF-E.pdf>.

Authors' Addresses

Emile Stephan (editor)
Orange
2, avenue Pierre Marzin
22300 Lannion
France
Email: emile.stephan@orange.com

Roland Schott
Deutsche Telekom
Deutsche-Telekom-Allee 9
64295 Darmstadt
Germany
Email: Roland.Schott@telekom.de

Diego Lopez
Telefonica
Email: diego.r.lopez@telefonica.com

Xiaodong Duan
China Mobile
Email: duanxiaodong@chinamobile.com

Lionel Morand
Huawei
18 QUAI DU POINT DU JOUR
92100 BOULOGNE-BILLANCOURT
France
Email: lionel.morand@huawei.com