NOTIFICATION HANDLING FOR VERTICAL FEDERATED LEARNING ENABLEMENT

TECHNICAL FIELD

[0001] The present disclosure relates to wireless communications, and more specifically to vertical federated learning enablement.

BACKGROUND

[0002] A wireless communications system may include one or multiple network communication devices, such as base stations, which may support wireless communications for one or multiple user communication devices, which may be otherwise known as user equipment (UE), or other suitable terminology. The wireless communications system may support wireless communications with one or multiple user communication devices by utilizing resources of the wireless communication system (e.g., time resources (e.g., symbols, slots, subframes, frames, or the like) or frequency resources (e.g., subcarriers, carriers, or the like)). Additionally, the wireless communications system may support wireless communications across various radio access technologies including third generation (3G) radio access technology, fourth generation (4G) radio access technology, fifth generation (5G) radio access technology, among other suitable radio access technologies beyond 5G (e.g., sixth generation (6G)).

[0003] Use cases and requirements have been identified for the support of Artificial intelligence (AI)/ Machine Learning (ML) model distribution, transfer, and training for various applications (e.g., video/speech recognition, robot control, automotive control), and also covering support for Distributed AI training/inference based on direct device connection.

SUMMARY

[0004] An article “a” before an element is unrestricted and understood to refer to “at least one” of those elements or “one or more” of those elements. The terms “a,” “at least one,” “one or more,” and “at least one of one or more” may be interchangeable. As used herein, including in the claims, “or” as used in a list of items (e.g., a list of items prefaced by a phrase such as “at least one of” or “one or more of” or “one or both of”) indicates an

inclusive list such that, for example, a list of at least one of A, B, or C means A or B or C or AB or AC or BC or ABC (i.e., A and B and C). Also, as used herein, the phrase “based on” shall not be construed as a reference to a closed set of conditions. For example, an example step that is described as “based on condition A” may be based on both a condition A and a condition B without departing from the scope of the present disclosure. In other words, as used herein, the phrase “based on” shall be construed in the same manner as the phrase “based at least in part on”. Further, as used herein, including in the claims, a “set” may include one or more elements.

[0005] Some implementations of the method and apparatuses described herein may include a network entity for wireless communication with an AI/ML Service Registry, the AI/ML Service Registry being configured to manage a subscription for notifications related to a Federated Learning (FL) service, the network entity comprising: at least one memory; and at least one processor coupled with the at least one memory and configured to cause the network entity to: send a subscription request to the AI/ML Service Registry, wherein the subscription request comprises a request to receive information on at least one application layer event type; and receive, from the AI/ML Service Registry, information related to the at least one application layer event type.

[0006] The network entity may be further configured to receive a subscription confirmation notification.

[0007] The subscription confirmation notification may comprise a unique identifier for the subscription.

[0008] The application layer event type may be at least one of: an availability or unavailability of one or more candidate FL members; an availability change of a selected FL member; an FL service Application Programming Interface (API) update; an update on a computational load of a selected FL member; an update on an energy consumption of a selected FL member; an update on FL model monitoring information; a change in capability of an FL member; an update relating to the functioning of an FL service; or, an update relating to system load conditions.

[0009] The FL member may be one of: a vertical application layer (VAL) server; an application function; an application enabler entity; an edge enabler entity, an analytics enabler entity; a 5G network data analytics function (NWDAF); a management domain

analytics service (MDAS); an edge server, a cloud server, a VAL user equipment (UE); a UE; or, a UE application.

[0010] The subscription request may further comprise at least one of: identity information of the subscribing entity; a machine learning (ML) model ID; a ML model profile; a request for information on a target area; a requested time of validity for the subscription; an FL member type; an FL member ID; a list of UE IDs for which the subscription applies; a list of UE types for which the subscription applies; a PLMN ID; a service ID; service API criteria; notification reception information; ML model interoperability information; or ML vendor interoperability information.

[0011] The network entity may be one of: a vertical application layer (VAL) server; an application function; an application enabler entity; an edge enabler entity, an edge server, a cloud server, a northbound API invoker, an FL service consumer, an analytics enabler entity; a 5G network data analytics function (NWDAF); a management domain analytics service (MDAS); a VAL user equipment (UE); or, a UE application.

[0012] The processor may comprise at least one controller coupled with at least one memory and configured to cause the processor to send a subscription request to the AI/ML Service Registry, wherein the subscription request comprises a request to receive information on at least one application layer event type; and receive, from the AI/ML Service Registry, information related to the at least one application layer event type.

[0013] Some implementations of the method and apparatuses described herein may include an AI/ML Service Registry for managing subscriptions related to a Federated Learning (FL) service, the network entity comprising: at least one memory; and at least one processor coupled with the at least one memory and configured to cause the network entity to: obtain, from a subscriber entity, a subscription request, wherein the subscription request comprises a request to receive information on at least one application layer event type; authorize subscription of the subscriber entity; and notify the subscriber entity of information related to the at least one application layer event type.

[0014] The AI/ML Service Registry may be one of: a CAPIF core function, an edge platform repository, a cloud platform repository, an application layer analytics data repository function (A-ADRF), an ADRF, or a network repository function (NRF).

[0015] The AI/ML Service Registry may be further configured to send a subscription confirmation notification to the subscriber entity.

[0016] The authorizing may comprise at least one of: checking permissions of the subscriber entity to be notified of application layer events; registering the subscriber entity be notified of application layer events.

[0017] The authorizing may comprise at least one of: sending a request to an AI/ML Pipeline Manager to check permissions of the subscriber entity to be notified of application layer events; or, sending a request to the AI/ML Pipeline Manager to register the subscriber entity to be notified of application layer events.

[0018] The AI/ML Service Registry may be further configured to receive a trigger notification from at least one FL member indicating a change in the availability of the FL member, or a registration of the FL member to the FL service.

[0019] The AI/ML Service Registry may be further configured to generate information relating to an application layer event based on the trigger notification.

[0020] Some implementations of the methods and apparatuses described herein may further include a method for wireless communication between a subscriber entity and an AI/ML Service Registry, the AI/ML Service Registry being configured to manage subscription for notifications related to a Federated Learning (FL) service, the method comprising: sending, by the subscriber entity, a subscription request to the AI/ML Service Registry, wherein the subscription request comprises a request to receive information on at least one application layer event type; obtaining, by the AI/ML Service Registry, the subscription request; authorizing, by the AI/ML Service Registry, subscription of the subscriber entity; notifying the subscriber entity, by the AI/ML Service Registry, of information related to the at least one application layer event type and receiving at the subscriber entity, from the AI/ML Service Registry, information related to the at least one application layer event type.

BRIEF DESCRIPTION OF THE DRAWINGS

[0021] Figure 1 illustrates an example of a wireless communications system in accordance with aspects of the present disclosure.

[0022] Figure 2 illustrates an overview of Machine Learning (ML) model pipeline operations and placement in accordance with aspects of the present disclosure.

[0023] Figure 3 illustrates a high level Service Enabler Architecture Layer (SEAL)/ Application Data Analytics Enablement Service (ADAES) architecture in accordance with aspects of the present disclosure.

[0024] Figure 4 illustrates a coordinated ADAES deployment model in accordance with aspects of the present disclosure.

[0025] Figure 5 illustrates an example of a subscriber entity in accordance with aspects of the present disclosure.

[0026] Figure 6 illustrates an example of a processor in accordance with aspects of the present disclosure.

[0027] Figure 7 illustrates an example of an AI/ML service registry in accordance with aspects of the present disclosure.

[0028] Figure 8 illustrates a flow diagram of an event subscription method in accordance with aspects of the present disclosure.

[0029] Figure 9 illustrates a flow diagram of an event subscription method in accordance with aspects of the present disclosure.

[0030] Figure 10 illustrates a flow diagram of an event subscription method in accordance with aspects of the present disclosure.

[0031] Figure 11 illustrate a flowchart of a method performed by a subscriber entity in accordance with aspects of the present disclosure.

[0032] Figure 12 illustrate a flowchart of a method performed by an AI/ML service registry in accordance with aspects of the present disclosure.

DETAILED DESCRIPTION

[0033] 3GPP has identified use cases and requirements (in TS 22.261) for the support of Artificial intelligence (AI)/ Machine Learning (ML) model distribution, transfer, and training for various applications (e.g., video/speech recognition, robot control, automotive

control), and also covering support for Distributed AI training/inference based on direct device connection.

[0034] Relating to AI/ML support, the 3GPP SA2 (system architecture group) has investigated the following two aspects:

- Enhancements to a Network Data Analytics Function (NWDAF) for supporting Federated learning (FL) enabled analytics to improve an analytics output. In this enhancement, NWDAFs acting as FL members (servers or clients) register to a Network Repository Function (NRF) and can be discovered by other Network Functions (NFs) (currently only by other NWDAFs).
- Enhancements to 5G Core (5GC) (as specified in TS 23.501 clause 5.46) for assisting AI/ML (including FL) operations in an application layer (between one or more AI/ML users and an AI/ML server). In this enhancement, a Network Exposure Function (NEF) may assist the AI/ML application server in scheduling available UE(s) to participate in the AI/ML Federated Learning. Also, 5GC may assist in the selection of UEs to serve as FL clients, by providing a list of target member UE(s), to an Application Function (AF) that is subscribed to the NEF, to be notified about a subset list of UE(s) (i.e., a list of candidate UE(s) that fulfil certain filtering criteria).

[0035] In addition, 3GPP SA6 has specified an application layer architecture to enable data analytics as a new Service Enabler Architecture Layer (SEAL) service, also known as an Application Data Analytics Enablement Service (ADAES), in 3GPP TS 23.436. Such architecture provides an application layer analytics framework which offers generic analytics exposure and value-added services for verticals and ASPs. This includes application layer analytics related to end-to-end application performance, edge load, service API availability, location accuracy, and slice-related performance and fault analysis. In Rel-19, the ADAES functional architecture is expected to be enhanced to further improve functionality in 3GPP, in order to improve support for analytics using AI/ML methods.

[0036] One aspect not covered in these specifications is how to support the registration, discovery, notification handling, and subscriptions for entities which will serve as FL

members in the AI/ML model lifecycle (or pipeline) (FL clients, FL servers, FL aggregators in case of hierarchical FL, data collection coordinators).

[0037] An FL member can be equivalent to an FL functionality, FL client, FL server, FL aggregator, or a combination thereof. The FL member is an entity or logical functionality or capability or service of an entity, which undertakes some processing of the ML model pipeline operations, and in particular the ML model training and/or inference, for the scenario where the ML model pipeline operations employ a federated learning technique for the training/inference.

[0038] One example of FL member can be an FL client e.g. at the UE or edge server function, which is responsible for locally training the ML model using its own data set.

[0039] The AI/ML pipeline includes the following aspects as shown in Fig. 2, and can include blocks which can be distributed over different entities in the application layer as well as at the 5G network side. For example, in Fig. 2, the AI/ML model pipeline comprises: A data collection unit 202, a data preparation unit 204, a ML model selection unit 206, a Model training unit 208, a ML model deployment unit 210, a ML model inference unit 212, and a prediction/prescription unit 214. The AI/ML model pipeline can be equivalent to the AI/ML model lifecycle in certain implementations.

[0040] For the case of FL, the ML model pipeline includes more than one entity which is configured to operate in parallel with another entity to implement the training and/or inference blocks. This adds additional complexity to the ML model pipeline, as the ML model pipeline provides functionality to allow different entities to be registered and discovered as FL members, including diverse entities from different domains and possibly from different stakeholders.

[0041] In 3GPP, analytics have been provided from the 5GC perspective, an enablement layer perspective, as well as from OAM perspective.

[0042] NWDAF provides analytics services at the 5GC side and can expose the analytics to other NFs via Service Based Interface (SBI) Application Programming Interfaces (APIs) and to AFs via NEF or Common API Framework (CAPIF) APIs. Some of the analytics services which collect data from user-plane sessions (from UPF and/or AF) and are used for the application server to acquire predictions or statistics from NWDAF are the following:

1) Observed Service Experience related network data analytics include: Service experience for a UE or a group of UEs, or any UE in an Application or a set of Applications over a specific User Plane (UP) path (e.g. a User Plane Function (UPF), a Data Network Access Identifier (DNAI) and an Edge Computing (EC) server).

2) Wireless Local Area network (WLAN) performance analytics (UE Uplink downlink (UL/DL) data rate and traffic volume)

3) UE communication analytics (NWDAF collects UE UL/DL data rate and traffic volume from a UPF and N4 session level data)

4) User data congestion analytics

-NWDAF acquires UL/DL throughput and peak UL/DL throughput from a UPF or AF

5) Redundant Transmission Experience related analytics (UL/DL packet delay from UE to UPF.)

6) DN Performance Analytics (performance data from an AF, e.g. Average Packet Delay, Average Loss Rate and Throughput)

7) QoS sustainability analytics

[0043] In Rel-18, the NWDAF also includes FL enabled analytics. In particular, federated learning among multiple NWDAFs is a machine learning technique in a core network that trains an ML Model across multiple decentralized entities holding local data sets, without exchanging/sharing the local data sets. This approach stands in contrast to traditional centralized machine learning techniques where all the local datasets are uploaded to one server, which allows critical issues such as data privacy, data security, and data access rights to be addressed. For Federated Learning supported by multiple NWDAFs containing a Model Training Logical Function (MTLF), one NWDAF contains an MTLF acting as an FL server (called FL server NWDAF for short) and multiple NWDAFs contain an MTLF acting as an FL client (called FL client NWDAF for short). The FL server NWDAF discovers and selects FL client NWDAFs to participant in an FL procedure, requests FL client NWDAFs to perform local model training and to report local model information, generates a global ML model by aggregating local model information from the

FL client NWDAFs, and sends the global ML model back to the FL client NWDAFs and repeats the training iteration if needed.

[0044] Each FL client NWDAF locally trains an ML model that is provided by the FL server NWDAF with the available local data set, which includes the data that the FL client NWDAF is not allowed to share with others (due to e.g., data privacy, data security, data access rights); reports the trained local ML model information to the FL server NWDAF; and receives the global ML model feedback from the FL server NWDAF and repeats the training iteration if needed.

[0045] A Management data analytics service (MDAS) provides data analytics of different network management matters including, for example, load level and/or resource utilisation. For example, the MDAS for a network function (NF) can create a management job to collect the NF's load related performance data, e.g., resource usage status of the NF, using management interfaces, such as MnS. The analysis of the collected data can be per area, slice or slice subnet, and may provide forecast of resource usage information in a predefined future time. This analysis may also recommend appropriate actions, e.g., scaling of resources, admission control, load balancing of traffic, etc., or perform root cause analysis to identify a management problem, e.g., a fault or congestion, and recommend a solution.

[0046] An application data analytics enablement server or service (ADAES) is defined (SA6, 3GPP TS 23.436) for performing app layer and edge / cloud analytics outside the 3GPP domain. ADAES can be seen as an AF which has analytics capability, and also has an interface to the UE side (to an ADAE client) as well as to an Operations and Maintenance (OAM) function or service. Fig. 3 illustrates a high-level architecture for an ADAE service. In this figure, the VAL server(s) communicates with the ADAES over an ADAES reference point. The ADAES, acting as an AF, may communicate with the 5G Core Network functions (over an N33 reference point to the NEF and an N6 reference point to the UPF) and the OAM function or service (over an ADAE-OAM interface).

[0047] ADAES supports analytics (e.g., VAL server performance, edge load analytics, location analytics etc.) and it is also possible that such analytics methods can be ML-enabled. There are different deployments and business models for ADAES supported in TS 23.436. For example, ADAES can be deployed within a Public Land Mobile Network

(PLMN), or at an Edge Computing Service Provider (ECSP) or at vertical domain (e.g., automaker, factory owner, stadium service provider).Regarding the deployment models, there are three possible scenarios (centralized, distributed, and coordinated).

[0048] Figure 4 illustrates a coordinated ADAES deployment model. As stated at TS 23.436, multiple ADAESs can be located at different Data Networks (DNs) / Edge DNs (EDNs) and can be deployed by the same ADAE provider. Such coordinated deployments allow for local to global analytics derivation (which may be beneficial in improving the analytics confidence level). The centrally deployed ADAES can also act as an ADAE analytics aggregator entity and can configures the edge deployed ADAES to derive analytics on different sub-areas. One example is the use of analytics for an EDN#1 or EDN#2 load, which helps in predicting the VAL server performance at a centrally located ADAES. Such deployment is also applicable for ML-based analytics methods, like supervised learning, where the centrally located ADAES acts as an ML model training entity, and the edge located ADAESs can act as an ML model inference entity (using edge data to improve the prediction accuracy).

[0049] The statistics/predictions of the edge deployed ADAES correspond to the ADAES service areas, which are equivalent to the EES/EAS service areas. The central ADAE server covers all PLMN areas and is used to coordinate or jointly perform analytics with the distributed ADAES. Such analytics services can be provided to consumers at the central DN, like the VAL servers or SEAL services or even at the PLMN side (e.g., NWDAF consuming service experience analytics).

[0050] Current ADAES offers AI/ML enabled services, however one of the key aspects which are planned to be covered in Rel-19 is to study the implications of new AI/ML methods for ADAES analytics. In particular, to study enhanced or new enablement capabilities for supporting ML model training / inference and federated learning in coordinated ADAES deployments. This may support for example the assistance for the discovery and selection of the SA6 enablers (e.g. ADAEC, edge or central ADAES) to serve as ML training/inference entities or FL members in the federated learning case.

[0051] In 3GPP, the exposure of APIs to 3rd parties can be performed via NEF/CAPIF or via EGMF or via enablement APIs.

[0052] In SA2, some of the NEF APIs related to the exposure to verticals/ ASPs include:

Service Name	Description
Nnef_EventExposure	Provides support for event exposure.
Nnef_ParameterProvision	Provides support to provision information which can be used for the UE in 5GS.
Nnef_Trigger	Provides support for device triggering.
Nnef_TrafficInfluence	Provide the ability to influence traffic routing.
Nnef_AFsessionWithQoS	Requests the network to provide a specific QoS for an AS session.
Nnef_ServiceParameter	Provides support to provision service specific information.
Nnef_APISupportCapability	Provides support for awareness on availability or expected level of a service API.
Nnef_AnalyticsExposure	Provides support for exposure of network analytics.
Nnef_Location	Provides the capability to deliver UE location to AF.
Nnef_AMInfluence	Provides the ability to influence access and mobility management related policies for one or multiple UEs.
Nnef_EASDeployment	EAS deployment service.
Nnef_UEId	UE Identifier service, which supports to retrieve AF specific UE Identifier based on UE address

[0053] Regarding AI/ML related exposure, the AF may subscribe to NEF monitoring events to assist its application AI/ML operations and it can also use the NEF analytics exposure. For example, the AF may subscribe to a session inactivity time monitoring event to assist the AI/ML application server in scheduling available UE(s) to participate in the AI/ML operation (e.g., Federated Learning) or can subscribe to network performance analytics in the area where candidate UEs are residing. In addition, the AF may subscribe to NEF to be notified on the traffic volume exchanged between the UE and the AI/ML application server to assist the AF with the transfer of AI/ML data.

[0054] A list of APIs can be provided to 3rd party application servers (e.g VAL servers or Edge Application Servers). APIs are defined for different types of enablement services, as follows:

- EDGEAPP (TS 23.558)
- SEAL (TS 23.434)
- V2XAPP (TS 23.286), UASAPP (TS23.255),

[0055] In SA6 [3GPP TS 23.222, 3GPP TS 29.222], a Common API Framework (CAPIF) was developed to enable a unified Northbound API framework across 3GPP network functions, and to ensure that there is a single and harmonized approach for API development. Some key functionalities in CAPIF are:

- CAPIF Core Function (CCF) is a repository of all, PLMN and 3rd party, service APIs.
- API Exposing Function (AEF) is the provider of the services as APIs.
- API Invoker is typically the applications that require service from the service providers.
- API management function enables the API provider to perform administration of the service APIs. This includes auditing the service API invocation logs received from the CAPIF core function, Monitoring the events reported by the CAPIF core function etc.
- API publishing function enables the API provider to publish the service APIs information in order to enable the discovery of service APIs by the API invoker.

[0056] The CAPIF is hosted within the PLMN operator network or trust domain. The API invoker is typically provided by a 3rd party application provider who has service agreement with PLMN operator. The API invoker may reside within the same trust domain as the PLMN operator network.

[0057] In a reference point based model, the API invoker within the PLMN trust domain interacts with the CAPIF via CAPIF-1 and CAPIF-2. The API invoker from outside the PLMN trust domain interacts with the CAPIF via CAPIF-1e and CAPIF-2e. The API exposing function, the API publishing function and the API management function of the API provider domain (together known as API provider domain functions) within the PLMN

trust domain interacts with the CAPIF core function via CAPIF-3, CAPIF-4 and CAPIF-5 respectively.

[0058] There is a need to provide a system which allows external consumers/subscribers or involved entities of the AI/ML services to be made aware of and receive notification on the availability and capabilities of the registered FL members (which can be heterogeneous and may belong to different domains) that can be potentially used for the AI/ML model training and/or inference. Here heterogeneous refers to both external and network entities.

[0059] Embodiments of the invention relate to the sending of a subscription request, from a subscriber entity, to an AI/ML Service Registry, the subscription request comprising a request to receive information on at least one application layer event type. The AI/ML service registry sends, to the subscriber entity, a notification comprising information related to the at least one application layer event type. This advantageously allows for a subscriber client to be made aware of changes in FL members' changes of status without the FL needing to be registered to the same domain.

[0060] Aspects of the present disclosure are described in the context of a wireless communications system.

[0061] Figure 1 illustrates an example of a wireless communications system 100 in accordance with aspects of the present disclosure. The wireless communications system 100 may include one or more NE 102, one or more UE 104, and a core network (CN) 106. The wireless communications system 100 may support various radio access technologies. In some implementations, the wireless communications system 100 may be a 4G network, such as an LTE network or an LTE-Advanced (LTE-A) network. In some other implementations, the wireless communications system 100 may be a NR network, such as a 5G network, a 5G-Advanced (5G-A) network, or a 5G ultrawideband (5G-UWB) network. In other implementations, the wireless communications system 100 may be a combination of a 4G network and a 5G network, or other suitable radio access technology including Institute of Electrical and Electronics Engineers (IEEE) 802.11 (Wi-Fi), IEEE 802.16 (WiMAX), IEEE 802.20. The wireless communications system 100 may support radio access technologies beyond 5G, for example, 6G. Additionally, the wireless

communications system 100 may support technologies, such as time division multiple access (TDMA), frequency division multiple access (FDMA), or code division multiple access (CDMA), etc.

[0062] The one or more NE 102 may be dispersed throughout a geographic region to form the wireless communications system 100. One or more of the NE 102 described herein may be or include or may be referred to as a network node, a base station, a network element, a network function, a network entity, a radio access network (RAN), a NodeB, an eNodeB (eNB), a next-generation NodeB (gNB), or other suitable terminology. An NE 102 and a UE 104 may communicate via a communication link, which may be a wireless or wired connection. For example, an NE 102 and a UE 104 may perform wireless communication (e.g., receive signaling, transmit signaling) over a Uu interface.

[0063] An NE 102 may provide a geographic coverage area for which the NE 102 may support services for one or more UEs 104 within the geographic coverage area. For example, an NE 102 and a UE 104 may support wireless communication of signals related to services (e.g., voice, video, packet data, messaging, broadcast, etc.) according to one or multiple radio access technologies. In some implementations, an NE 102 may be moveable, for example, a satellite associated with a non-terrestrial network (NTN). In some implementations, different geographic coverage areas 112 associated with the same or different radio access technologies may overlap, but the different geographic coverage areas may be associated with different NE 102.

[0064] The one or more UE 104 may be dispersed throughout a geographic region of the wireless communications system 100. A UE 104 may include or may be referred to as a remote unit, a mobile device, a wireless device, a remote device, a subscriber device, a transmitter device, a receiver device, or some other suitable terminology. In some implementations, the UE 104 may be referred to as a unit, a station, a terminal, or a client, among other examples. Additionally, or alternatively, the UE 104 may be referred to as an Internet-of-Things (IoT) device, an Internet-of-Everything (IoE) device, or machine-type communication (MTC) device, among other examples.

[0065] A UE 104 may be able to support wireless communication directly with other UEs 104 over a communication link. For example, a UE 104 may support wireless

communication directly with another UE 104 over a device-to-device (D2D) communication link. In some implementations, such as vehicle-to-vehicle (V2V) deployments, vehicle-to-everything (V2X) deployments, or cellular-V2X deployments, the communication link 114 may be referred to as a sidelink. For example, a UE 104 may support wireless communication directly with another UE 104 over a PC5 interface.

[0066] An NE 102 may support communications with the CN 106, or with another NE 102, or both. For example, an NE 102 may interface with other NE 102 or the CN 106 through one or more backhaul links (e.g., S1, N2, N3, or network interface). In some implementations, the NE 102 may communicate with each other directly. In some other implementations, the NE 102 may communicate with each other or indirectly (e.g., via the CN 106. In some implementations, one or more NE 102 may include subcomponents, such as an access network entity, which may be an example of an access node controller (ANC). An ANC may communicate with the one or more UEs 104 through one or more other access network transmission entities, which may be referred to as a radio heads, smart radio heads, or transmission-reception points (TRPs).

[0067] The CN 106 may support user authentication, access authorization, tracking, connectivity, and other access, routing, or mobility functions. The CN 106 may be an evolved packet core (EPC), or a 5G core (5GC), which may include a control plane entity that manages access and mobility (e.g., a mobility management entity (MME), an access and mobility management functions (AMF)) and a user plane entity that routes packets or interconnects to external networks (e.g., a serving gateway (S-GW), a Packet Data Network (PDN) gateway (P-GW), or a user plane function (UPF)). In some implementations, the control plane entity may manage non-access stratum (NAS) functions, such as mobility, authentication, and bearer management (e.g., data bearers, signal bearers, etc.) for the one or more UEs 104 served by the one or more NE 102 associated with the CN 106.

[0068] The CN 106 may communicate with a packet data network over one or more backhaul links (e.g., via an S1, N2, N3, or another network interface). The packet data network may include an application server. In some implementations, one or more UEs 104 may communicate with the application server. A UE 104 may establish a session (e.g., a protocol data unit (PDU) session, or the like) with the CN 106 via an NE 102. The CN 106

may route traffic (e.g., control information, data, and the like) between the UE 104 and the application server using the established session (e.g., the established PDU session). The PDU session may be an example of a logical connection between the UE 104 and the CN 106 (e.g., one or more network functions of the CN 106).

[0069] In the wireless communications system 100, the NEs 102 and the UEs 104 may use resources of the wireless communications system 100 (e.g., time resources (e.g., symbols, slots, subframes, frames, or the like) or frequency resources (e.g., subcarriers, carriers)) to perform various operations (e.g., wireless communications). In some implementations, the NEs 102 and the UEs 104 may support different resource structures. For example, the NEs 102 and the UEs 104 may support different frame structures. In some implementations, such as in 4G, the NEs 102 and the UEs 104 may support a single frame structure. In some other implementations, such as in 5G and among other suitable radio access technologies, the NEs 102 and the UEs 104 may support various frame structures (i.e., multiple frame structures). The NEs 102 and the UEs 104 may support various frame structures based on one or more numerologies.

[0070] One or more numerologies may be supported in the wireless communications system 100, and a numerology may include a subcarrier spacing and a cyclic prefix. A first numerology (e.g., $\mu=0$) may be associated with a first subcarrier spacing (e.g., 15 kHz) and a normal cyclic prefix. In some implementations, the first numerology (e.g., $\mu=0$) associated with the first subcarrier spacing (e.g., 15 kHz) may utilize one slot per subframe. A second numerology (e.g., $\mu=1$) may be associated with a second subcarrier spacing (e.g., 30 kHz) and a normal cyclic prefix. A third numerology (e.g., $\mu=2$) may be associated with a third subcarrier spacing (e.g., 60 kHz) and a normal cyclic prefix or an extended cyclic prefix. A fourth numerology (e.g., $\mu=3$) may be associated with a fourth subcarrier spacing (e.g., 120 kHz) and a normal cyclic prefix. A fifth numerology (e.g., $\mu=4$) may be associated with a fifth subcarrier spacing (e.g., 240 kHz) and a normal cyclic prefix.

[0071] A time interval of a resource (e.g., a communication resource) may be organized according to frames (also referred to as radio frames). Each frame may have a duration, for example, a 10 millisecond (ms) duration. In some implementations, each frame may include multiple subframes. For example, each frame may include 10 subframes, and each

subframe may have a duration, for example, a 1 ms duration. In some implementations, each frame may have the same duration. In some implementations, each subframe of a frame may have the same duration.

[0072] Additionally or alternatively, a time interval of a resource (e.g., a communication resource) may be organized according to slots. For example, a subframe may include a number (e.g., quantity) of slots. The number of slots in each subframe may also depend on the one or more numerologies supported in the wireless communications system 100. For instance, the first, second, third, fourth, and fifth numerologies (i.e., $\mu=0$, $\mu=1$, $\mu=2$, $\mu=3$, $\mu=4$) associated with respective subcarrier spacings of 15 kHz, 30 kHz, 60 kHz, 120 kHz, and 240 kHz may utilize a single slot per subframe, two slots per subframe, four slots per subframe, eight slots per subframe, and 16 slots per subframe, respectively. Each slot may include a number (e.g., quantity) of symbols (e.g., OFDM symbols). In some implementations, the number (e.g., quantity) of slots for a subframe may depend on a numerology. For a normal cyclic prefix, a slot may include 14 symbols. For an extended cyclic prefix (e.g., applicable for 60 kHz subcarrier spacing), a slot may include 12 symbols. The relationship between the number of symbols per slot, the number of slots per subframe, and the number of slots per frame for a normal cyclic prefix and an extended cyclic prefix may depend on a numerology. It should be understood that reference to a first numerology (e.g., $\mu=0$) associated with a first subcarrier spacing (e.g., 15 kHz) may be used interchangeably between subframes and slots.

[0073] In the wireless communications system 100, an electromagnetic (EM) spectrum may be split, based on frequency or wavelength, into various classes, frequency bands, frequency channels, etc. By way of example, the wireless communications system 100 may support one or multiple operating frequency bands, such as frequency range designations FR1 (410 MHz – 7.125 GHz), FR2 (24.25 GHz – 52.6 GHz), FR3 (7.125 GHz – 24.25 GHz), FR4 (52.6 GHz – 114.25 GHz), FR4a or FR4-1 (52.6 GHz – 71 GHz), and FR5 (114.25 GHz – 300 GHz). In some implementations, the NEs 102 and the UEs 104 may perform wireless communications over one or more of the operating frequency bands. In some implementations, FR1 may be used by the NEs 102 and the UEs 104, among other equipment or devices for cellular communications traffic (e.g., control information, data).

In some implementations, FR2 may be used by the NEs 102 and the UEs 104, among other equipment or devices for short-range, high data rate capabilities.

[0074] FR1 may be associated with one or multiple numerologies (e.g., at least three numerologies). For example, FR1 may be associated with a first numerology (e.g., μ=0), which includes 15 kHz subcarrier spacing; a second numerology (e.g., μ=1), which includes 30 kHz subcarrier spacing; and a third numerology (e.g., μ=2), which includes 60 kHz subcarrier spacing. FR2 may be associated with one or multiple numerologies (e.g., at least 2 numerologies). For example, FR2 may be associated with a third numerology (e.g., μ=2), which includes 60 kHz subcarrier spacing; and a fourth numerology (e.g., μ=3), which includes 120 kHz subcarrier spacing.

[0075] Figure 5 illustrates an example of a subscriber entity 500 in accordance with aspects of the present disclosure. The subscriber entity 500 may include a processor 502, a memory 504, a controller 506, and a transceiver 508. The processor 502, the memory 504, the controller 506, or the transceiver 508, or various combinations thereof or various components thereof may be examples of means for performing various aspects of the present disclosure as described herein. These components may be coupled (e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces.

[0076] The processor 502, the memory 504, the controller 506, or the transceiver 508, or various combinations or components thereof may be implemented in hardware (e.g., circuitry). The hardware may include a processor, a digital signal processor (DSP), an application-specific integrated circuit (ASIC), or other programmable logic device, or any combination thereof configured as or otherwise supporting a means for performing the functions described in the present disclosure.

[0077] The processor 502 may include an intelligent hardware device (e.g., a general-purpose processor, a DSP, a CPU, an ASIC, an FPGA, or any combination thereof). In some implementations, the processor 502 may be configured to operate the memory 504. In some other implementations, the memory 504 may be integrated into the processor 502. The processor 502 may be configured to execute computer-readable instructions stored in

the memory 504 to cause the subscriber entity 500 to perform various functions of the present disclosure.

[0078] The memory 504 may include volatile or non-volatile memory. The memory 504 may store computer-readable, computer-executable code including instructions when executed by the processor 502 cause the subscriber entity 500 to perform various functions described herein. The code may be stored in a non-transitory computer-readable medium such as the memory 504 or another type of memory. Computer-readable media includes both non-transitory computer storage media and communication media including any medium that facilitates transfer of a computer program from one place to another. A non-transitory storage medium may be any available medium that may be accessed by a general-purpose or special-purpose computer.

[0079] In some implementations, the processor 502 and the memory 504 coupled with the processor 502 may be configured to cause the subscriber entity 500 to perform one or more of the functions described herein (e.g., executing, by the processor 502, instructions stored in the memory 504). For example, the processor 502 may support wireless communication at the subscriber entity 500 in accordance with examples as disclosed herein. The subscriber entity 500 may be configured to support a means for sending a subscription request to an AI/ML Service Registry, wherein the subscription request comprises a request to receive information on at least one application layer event type; and receiving, from the AI/ML Service Registry, information related to the at least one application layer event type.

[0080] The controller 506 may manage input and output signals for the subscriber entity 500. The controller 506 may also manage peripherals not integrated into the subscriber entity 500. In some implementations, the controller 506 may utilize an operating system such as iOS®, ANDROID®, WINDOWS®, or other operating systems. In some implementations, the controller 506 may be implemented as part of the processor 502.

[0081] In some implementations, the subscriber entity 500 may include at least one transceiver 508. In some other implementations, the subscriber entity 500 may have more than one transceiver 508. The transceiver 508 may represent a wireless transceiver. The

transceiver 508 may include one or more receiver chains 510, one or more transmitter chains 512, or a combination thereof.

[0082] A receiver chain 510 may be configured to receive signals (e.g., control information, data, packets) over a wireless medium. For example, the receiver chain 510 may include one or more antennas for receiving the signal over the air or wireless medium. The receiver chain 510 may include at least one amplifier (e.g., a low-noise amplifier (LNA)) configured to amplify the received signal. The receiver chain 510 may include at least one demodulator configured to demodulate the received signal and obtain the transmitted data by reversing the modulation technique applied during transmission of the signal. The receiver chain 510 may include at least one decoder for decoding the demodulated signal to receive the transmitted data.

[0083] A transmitter chain 512 may be configured to generate and transmit signals (e.g., control information, data, packets). The transmitter chain 512 may include at least one modulator for modulating data onto a carrier signal, preparing the signal for transmission over a wireless medium. The at least one modulator may be configured to support one or more techniques such as amplitude modulation (AM), frequency modulation (FM), or digital modulation schemes like phase-shift keying (PSK) or quadrature amplitude modulation (QAM). The transmitter chain 512 may also include at least one power amplifier configured to amplify the modulated signal to an appropriate power level suitable for transmission over the wireless medium. The transmitter chain 512 may also include one or more antennas for transmitting the amplified signal into the air or wireless medium.

[0084] Figure 6 illustrates an example of a processor 600 in accordance with aspects of the present disclosure. The processor 600 may be an example of a processor configured to perform various operations in accordance with examples as described herein. The processor 600 may include a controller 602 configured to perform various operations in accordance with examples as described herein. The processor 600 may optionally include at least one memory 604, which may be, for example, an L1/L2/L3 cache. Additionally, or alternatively, the processor 600 may optionally include one or more arithmetic-logic units (ALUs) 606. One or more of these components may be in electronic communication or

otherwise coupled (e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces (e.g., buses).

[0085] The processor 600 may be a processor chipset and include a protocol stack (e.g., a software stack) executed by the processor chipset to perform various operations (e.g., receiving, obtaining, retrieving, transmitting, outputting, forwarding, storing, determining, identifying, accessing, writing, reading) in accordance with examples as described herein. The processor chipset may include one or more cores, one or more caches (e.g., memory local to or included in the processor chipset (e.g., the processor 600) or other memory (e.g., random access memory (RAM), read-only memory (ROM), dynamic RAM (DRAM), synchronous dynamic RAM (SDRAM), static RAM (SRAM), ferroelectric RAM (FeRAM), magnetic RAM (MRAM), resistive RAM (RRAM), flash memory, phase change memory (PCM), and others)).

[0086] The controller 602 may be configured to manage and coordinate various operations (e.g., signaling, receiving, obtaining, retrieving, transmitting, outputting, forwarding, storing, determining, identifying, accessing, writing, reading) of the processor 600 to cause the processor 600 to support various operations in accordance with examples as described herein. For example, the controller 602 may operate as a control unit of the processor 600, generating control signals that manage the operation of various components of the processor 600. These control signals include enabling or disabling functional units, selecting data paths, initiating memory access, and coordinating timing of operations.

[0087] The controller 602 may be configured to fetch (e.g., obtain, retrieve, receive) instructions from the memory 604 and determine subsequent instruction(s) to be executed to cause the processor 600 to support various operations in accordance with examples as described herein. The controller 602 may be configured to track memory address of instructions associated with the memory 604. The controller 602 may be configured to decode instructions to determine the operation to be performed and the operands involved. For example, the controller 602 may be configured to interpret the instruction and determine control signals to be output to other components of the processor 600 to cause the processor 600 to support various operations in accordance with examples as described herein. Additionally, or alternatively, the controller 602 may be configured to manage flow

of data within the processor 600. The controller 602 may be configured to control transfer of data between registers, arithmetic logic units (ALUs), and other functional units of the processor 600.

[0088] The memory 604 may include one or more caches (e.g., memory local to or included in the processor 600 or other memory, such RAM, ROM, DRAM, SDRAM, SRAM, MRAM, flash memory, etc.). In some implementations, the memory 604 may reside within or on a processor chipset (e.g., local to the processor 600). In some other implementations, the memory 604 may reside external to the processor chipset (e.g., remote to the processor 600).

[0089] The memory 604 may store computer-readable, computer-executable code including instructions that, when executed by the processor 600, cause the processor 600 to perform various functions described herein. The code may be stored in a non-transitory computer-readable medium such as system memory or another type of memory. The controller 602 and/or the processor 600 may be configured to execute computer-readable instructions stored in the memory 604 to cause the processor 600 to perform various functions. For example, the processor 600 and/or the controller 602 may be coupled with or to the memory 604, the processor 600, the controller 602, and the memory 604 may be configured to perform various functions described herein. In some examples, the processor 600 may include multiple processors and the memory 604 may include multiple memories. One or more of the multiple processors may be coupled with one or more of the multiple memories, which may, individually or collectively, be configured to perform various functions herein.

[0090] The one or more ALUs 606 may be configured to support various operations in accordance with examples as described herein. In some implementations, the one or more ALUs 606 may reside within or on a processor chipset (e.g., the processor 600). In some other implementations, the one or more ALUs 606 may reside external to the processor chipset (e.g., the processor 600). One or more ALUs 606 may perform one or more computations such as addition, subtraction, multiplication, and division on data. For example, one or more ALUs 606 may receive input operands and an operation code, which determines an operation to be executed. One or more ALUs 606 be configured with a

variety of logical and arithmetic circuits, including adders, subtractors, shifters, and logic gates, to process and manipulate the data according to the operation. Additionally, or alternatively, the one or more ALUs 606 may support logical operations such as AND, OR, exclusive-OR (XOR), not-OR (NOR), and not-AND (NAND), enabling the one or more ALUs 606 to handle conditional operations, comparisons, and bitwise operations.
The processor 600 may support wireless communication in accordance with examples as disclosed herein. The processor 600 may be configured to or operable to support a means for outputting a subscription request to the AI/ML Service Registry, wherein the subscription request comprises a request to receive information on at least one application layer event type; and obtaining, from the AI/ML Service Registry, information related to the at least one application layer event type.

[0091] Figure 7 illustrates an example of an AI/ML service registry 700 in accordance with aspects of the present disclosure. The AI/ML service registry 700 may include a processor 702, a memory 704, a controller 706, and a transceiver 708. The processor 702, the memory 704, the controller 706, or the transceiver 708, or various combinations thereof or various components thereof may be examples of means for performing various aspects of the present disclosure as described herein. These components may be coupled (e.g., operatively, communicatively, functionally, electronically, electrically) via one or more interfaces.

[0092] The AI/ML Service Registry may be one of: a CAPIF core function as defined in 3GPP TS 23.222 and TS 29.222, an edge platform repository, a cloud platform repository, an application layer analytics data repository function (A-ADRF) as defined in 3GPP TS 23.436, an ADRF as defined in 3GPP TS 23.288, or a network repository function (NRF).

[0093] The processor 702, the memory 704, the controller 706, or the transceiver 708, or various combinations or components thereof may be implemented in hardware (e.g., circuitry). The hardware may include a processor, a digital signal processor (DSP), an application-specific integrated circuit (ASIC), or other programmable logic device, or any combination thereof configured as or otherwise supporting a means for performing the functions described in the present disclosure.

[0094] The processor 702 may include an intelligent hardware device (e.g., a general-purpose processor, a DSP, a CPU, an ASIC, an FPGA, or any combination thereof). In some implementations, the processor 702 may be configured to operate the memory 704. In some other implementations, the memory 704 may be integrated into the processor 702. The processor 702 may be configured to execute computer-readable instructions stored in the memory 704 to cause the AI/ML service registry 700 to perform various functions of the present disclosure.

[0095] The memory 704 may include volatile or non-volatile memory. The memory 704 may store computer-readable, computer-executable code including instructions when executed by the processor 702 cause the AI/ML service registry 700 to perform various functions described herein. The code may be stored in a non-transitory computer-readable medium such as the memory 704 or another type of memory. Computer-readable media includes both non-transitory computer storage media and communication media including any medium that facilitates transfer of a computer program from one place to another. A non-transitory storage medium may be any available medium that may be accessed by a general-purpose or special-purpose computer.

In some implementations, the processor 702 and the memory 704 coupled with the processor 702 may be configured to cause the AI/ML service registry 700 to perform one or more of the functions described herein (e.g., executing, by the processor 702, instructions stored in the memory 704). For example, the processor 702 may support wireless communication at the AI/ML service registry 700 in accordance with examples as disclosed herein. The AI/ML service registry 700 may be configured to support a means for obtaining, from a subscriber entity, a subscription request, wherein the subscription request comprises a request to receive information on at least one application layer event type; authorize subscription of the subscriber entity; and notify the subscriber entity of information related to the at least one application layer event type. The AI/ML Service Registry 700 may be configured to support a means for obtaining a subscription request, the subscription request comprising a request to receive information on at least one application layer event type.

[0096] The controller 706 may manage input and output signals for the AI/ML service registry 700. The controller 706 may also manage peripherals not integrated into the AI/ML service registry 700. In some implementations, the controller 706 may utilize an operating system such as iOS®, ANDROID®, WINDOWS®, or other operating systems. In some implementations, the controller 706 may be implemented as part of the processor 702.

[0097] In some implementations, the AI/ML service registry 700 may include at least one transceiver 708. In some other implementations, the AI/ML service registry 700 may have more than one transceiver 708. The transceiver 708 may represent a wireless transceiver. The transceiver 708 may include one or more receiver chains 710, one or more transmitter chains 712, or a combination thereof.

[0098] A receiver chain 710 may be configured to receive signals (e.g., control information, data, packets) over a wireless medium. For example, the receiver chain 710 may include one or more antennas for receiving the signal over the air or wireless medium. The receiver chain 710 may include at least one amplifier (e.g., a low-noise amplifier (LNA)) configured to amplify the received signal. The receiver chain 710 may include at least one demodulator configured to demodulate the received signal and obtain the transmitted data by reversing the modulation technique applied during transmission of the signal. The receiver chain 710 may include at least one decoder for decoding the demodulated signal to receive the transmitted data.

[0099] A transmitter chain 712 may be configured to generate and transmit signals (e.g., control information, data, packets). The transmitter chain 712 may include at least one modulator for modulating data onto a carrier signal, preparing the signal for transmission over a wireless medium. The at least one modulator may be configured to support one or more techniques such as amplitude modulation (AM), frequency modulation (FM), or digital modulation schemes like phase-shift keying (PSK) or quadrature amplitude modulation (QAM). The transmitter chain 712 may also include at least one power amplifier configured to amplify the modulated signal to an appropriate power level suitable for transmission over the wireless medium. The transmitter chain 712 may also include one or more antennas for transmitting the amplified signal into the air or wireless medium.

[0100] Embodiments of the invention relate to allowing a subscriber/consumer (which can be, for example, a coordination server FL or an API invoker entity) to subscribe for receiving AI/ML model related events and to be notified on changes on the availability of the FL members which are to be used for ML model training.

[0101] In this procedure a subscribing entity/consumer initially subscribes to an AI/ML Registry (AMLR) for notification of events related to the availability and status of ML model members. This includes the subscription request, authorization by the AMLR, and a response from the AMLR regarding the outcome of the subscription request (positive or negative). Figure 8 is a flow diagram illustrating this process.

[0102] In step 801, the subscriber entity sends a subscription request to the AI/ML service registry. In step 802, the AI/ML service registry authorizes the request. In step 803, the AI/ML service registry sends a notification of the result of the subscription request to the subscriber entity. In step 804, a new FL member registers to the FL service. In step 805, the AI/ML service registry notifies the subscriber entity that a new FL member has registered to the FL service. An FL service may be a support service/capability or functionality which can be provided by the network operator (or in general the wireless communication system) to assist the FL process related to the ML model. Such assistance can be the training of the ML model by a wireless communication system entity (acting as an FL client), and providing the trained model to the ML model owner/provider or the 3rd party FL server.

[0103] The FL service can support different scenarios. An FL service can be deployed for one or more of the following: for a service profile, an application type, an application service, a service area, a group of UEs, a VAL server, a UE, an ML model identifier/profile, an edge data network, or an analytics event.

[0104] The subscriber entity can be for example an application server or client, an Application Function (AF), a Network Function (NF) or Management Function (MF) (e.g., an analytics coordination entity), or an API invoker entity which requires to know the available ML model members (for training and FL of the data model). Such subscription can be per analytics event or per ML model ID or profile, e.g., to be aware of the entities that serve as FL clients/servers for a given ML model or for a given analytics ID.

[0105] An event may be captured at AMLR in the following ways:

- a trigger is provided by the FL member to the AMLR to indicate a change in the availability of the FL member;
- a trigger is created internally at the AMLR based on monitoring of the API logs (e.g., via using CAPIF) for the APIs being used by the FL members (e.g., northbound APIs providing trained ML models);
- an event is created after translating an expected or a predicted change of capabilities or availability of the FL member (which is registered at the AMLR).

[0106] The AMLR notifies all subscribers /consumers on the new event or advertises/publishes the new information. Existing subscribers are entities (e.g., app servers) which already have subscribed for receiving notifications on the availability of FL members. Involved entities are the existing or candidate ML model members (ADAES, EDGEAPP, AI Enabler) which may use the information published by the AMLR and may belong to the same provider.

[0107] An aspect of the present disclosure relates to subscription, by a subscriber entity, for events related to FL member availability. This includes the subscription for events and event notifications, as well as a list of FL related events.

[0108] In order to be subscribed to receive notification of events, the subscribing entity is authorized to subscribe for the FL-related events. The method is described with reference to Figure 9. In step 901, the subscribing entity sends an event subscription request to the AI/ML Service Registry in order to receive notification of events related to the AI/ML model pipeline changes and in particular the availability of FL member for a target service or area or UEs.

[0109] If the AI/ML Service Registry is not a CCF, then the Subscriber entity is assumed to be aware of the AI/ML Service Registry ID/address (e.g., via CAPIF or pre-configuration or based on the SLA).

[0110] The FL-related event subscription request may include one or more of the following parameters:

Information element	**Description**
Identity information	Information to determine the identity of the subscribing entity

ML model ID or profile	The identity of the ML model for which the subscription is requested. Alternatively, a profile of the desired ML model can be requested
Target area information	The target service area for which the subscription applies. This area can be in form of topological area (list of cells/TAs) or set of DNAI/DNN or geographical coordinates for the coverage of the subscription.
Time of validity	The time validity for the subscription.
FL member type	The type of FL member which can be an FL server or FL client.
FL member ID	The ID of the entity serving as FL member (if known). This can be used for events where the FL member is known and possibly selected, and the subscriber entity wants to check the availability.
List of UE IDs /type	The list of UEs (and types) for which the subscription for event applies
PLMN ID	The PLMN for which the event subscription will apply (in the case of multi-PLMN support by the AIML service registry)
Service ID	The application service identity or profile for which the event subscription applies.
Service API criteria	The API related criteria include information relating to the service API status of the API which is used at the FL/ML model pipeline. This can be a failure of an API invocation event, a new API available event, an API version change event, an API location change event, etc., and other query information, for example, a service API identifier, a service API name, etc.
Event type	The event type requested, based on the List of FL related events (shown in table below).
Notification reception information	The information of the subscribing entity for receiving the notifications for the event.
ML model / vendor interoperability information	The ML model or vendor interoperability indication.

[0111] In step 902, upon receiving the event subscription request from the subscribing entity, the AI/ML Service Registry checks for the relevant authorization for the event subscription. This checking can be either local (based on information regarding the permissions of the subscriber to be notified about events related to the FL member availability), or there can be additional interaction between the AI/ML Service Registry and an AI/ML Pipeline Manager (which can be at an OAM or 3rd party enablement layer (e.g. at the edge/cloud platform of the AI/ML service provider)). The AI/ML Pipeline Manager may be responsible for supporting the provisioning and authorizing of any changes regarding the AI/ML pipeline operations. The AI/ML Pipeline manager in certain embodiments can be a capability of the API Management Function. The AI/ML Pipeline manager in certain implementations can be an orchestrator entity which configures, orchestrates and monitors the AI/ML pipeline and/or lifecycle per Federated Learning (FL) service or per application type or per vertical.

[0112] In step 903, if the authorization is successful, the AI/ML Service Registry stores the subscription information.

[0113] In step 904, the CAPIF core function sends an event subscription response indicating successful operation.

[0114] Examples of the event subscription response are provided below:

Information element	Description
Result	Indicates the success or failure of the event subscription operation
Subscription identifier	The unique identifier for the event subscription.

[0115] A list of FL-related events which the subscriber entity can subscribe for notification of is provided below:

Events	Events Description
Availability/Unavailability of candidate FL member(s)	The event type relates to an availability or unavailability indication of a candidate FL member (or a set of members), which can be a registered entity such as a candidate FL server or FL client. Such availability can be for a given service area, or UE or application type/profile, or for a given ML model ID/profile.
Availability changes of selected FL member	The event type relates to an availability change of a selected FL member, which can be a selected entity as a FL server or FL client. Such an availability change can be for a given service area, or UE or application type/profile, or for a given ML model ID/profile.
FL service API update	The event relates to an API update related to the FL service. This can be due to the relocation of an FL member to a different DN/EDN, and the need to provide an updated northbound API.
Monitoring FL member load	This event relates to the monitoring of the computational load for the requested FL member. Such a load can be estimated (based on measurements) or expected/predicted. This event may be triggered upon reaching a threshold for high load, can be requested on-demand, or can be the result of periodical monitoring.
Monitoring FL member energy	This event relates to the monitoring of the energy consumption of the requested FL member. The energy consumption can be estimated (based on measurements) or expected/predicted. This event may be triggered upon reaching a threshold for high energy usage, can be requested on-demand, or can be the result of periodical monitoring..

FL model monitoring info	This event relates to a time schedule and latency for the FL training. A latency measure is determined with respect to a target latency, i.e., when the FL model training shall be completed. This event may also relate to an affected area of interest where FL members are residing or covering.
FL member capability change	An FL member may change its capability. Especially for UE or AF, the capability for performing some FL related tasks (e.g., training, reporting, data collection) may be changed due to the energy / load, processing capabilities change, connectivity change / congestion/interference.
System related events	Alarm events providing fault information for the FL service
Performance related events	Events related to system load conditions

[0116] Once the subscriber entity is subscribed to the AI/ML service registry, the following steps may be performed as shown in Figure 10.

[0117] In step 1001, the AI/ML Service Registry receives a trigger notification from a Candidate or Selected FL member, which has already registered. This trigger can be the registration of a new FL member at the AI/ML Service registry which may act as a trigger in this procedure. A trigger notification message includes one or more of the following parameters:

Information element	**Description**
ML model ID or profile	The identity of the ML model for which the subscription is needed.
FL member ID	The ID of the entity serving as an FL member (if known).
List of UE IDs	The list of UEs (and types) for which the trigger applies
Service ID	The application service identity or profile for which the trigger applies.

Trigger report	A trigger report comprises a notification on a change of availability or of the capabilities related to the FL member. This report may also include measurements or notifications (based on triggers) of a change in the load/energy status for the FL member.
API change indication	The expected change of service API endpoint for the FL member (e.g., due to FL member mobility to different service area).
Cause	The cause for the trigger notification relating to an unavailability of an FL member
Predicted timer	Optionally, a time window in which the FL member is going to be unavailable or available in the future (e.g., in case of predicted mobility/unavailability).
ML model / vendor interoperability information	An ML model or vendor interoperability indication.

[0118] In step 1002, the AI/ML Service Registry generates events to be consumed by the subscribing entity(s), based on the trigger notification or based on the received registration from a new FL member.

[0119] In step 1003, for the generated event, the AI/ML Service Registry retrieves the list of corresponding subscriptions.

[0120] In step 1004, the AI/ML Service Registry sends FL-related event notifications to all the subscribing entity(s) that have subscribed for the event matching the criteria. If notification reception information is available as part of the subscribing entity event subscription, then the notification reception information is used by the AI/ML Service Registry to send event notifications to the subscribing entity. Such FL-related Event notifications may include one or more of the following parameters:

Information element	Description
Subscription identifier	The unique identifier of the event subscription
Event identifier	The unique identifier for the event. For the list of events, refer to the table with the list of FL related events.
Event related information	The event related information (e.g., time at which the event originated or is planned to be executed, location of event)
Event content	The content of the event information.
> List of FL member ID /address	A list of the FL members and their addresses
>> FL member enter / leave	An indication of an FL member entering or leaving the available list
>> FL member updated capability	An update on the capabilities of the FL members.
>> FL member load status	The load information if the event relates to load monitoring
>> FL member energy status	The energy information if the event relates to energy monitoring
>>Time availability of member	The available times for an FL member to perform part of an FL service
>> area of interest or location (predicted or expected)	An area of interest for an FL service to be performed
>> interoperability info	Information on interoperability of FL members.

[0121] In step 1005, the subscribing entity sends an event notification acknowledgement to the AI/ML Service Registry for the event notification received.

[0122] Figure 11 illustrates a flowchart of a method in accordance with aspects of the present disclosure. The operations of the method may be implemented by a subscriber entity as described herein. In some implementations, the subscriber entity may execute a set of instructions to control the function elements of the subscriber entity to perform the described functions.

[0123] At 1102, the method may include sending a subscription request to an AI/ML Service Registry, wherein the subscription request comprises a request to receive information on at least one application layer event type. The operations of 1102 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 1102 may be performed by a subscriber entity as described with reference to Figure 5.

[0124] At 1104, the method may include receiving, from the AI/ML Service Registry, information related to the at least one application layer event type from the AI/ML Service Registry, information related to the at least one application layer event type. The operations of 1104 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 1104 may be performed by a subscriber entity as described with reference to Figure 5.

[0125] It should be noted that the method described herein describes a possible implementation, and that the operations and the steps may be rearranged or otherwise modified and that other implementations are possible.

[0126] Figure 12 illustrates a flowchart of a method in accordance with aspects of the present disclosure. The operations of the method may be implemented by a AI/ML Service Registry as described herein. In some implementations, the AI/ML Service Registry may execute a set of instructions to control the function elements of the AI/ML Service Registry to perform the described functions.

[0127] At 1202, the method may include obtaining, from a subscriber entity, a subscription request, wherein the subscription request comprises a request to receive information on at least one application layer event type. The operations of 1202 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 1202 may be performed by an AI/ML Service Registry as described with reference to Figure 7.

[0128] At 1204, the method may include authorizing subscription of the subscriber entity. The operations of 1204 may be performed in accordance with examples as described

herein. In some implementations, aspects of the operations of 1204 may be performed by a AI/ML Service Registry as described with reference to Figure 7.

[0129] At 1206, the method may include notifying the subscriber entity of information related to the at least one application layer event type. The operations of 1206 may be performed in accordance with examples as described herein. In some implementations, aspects of the operations of 1206 may be performed by an AI/ML Service Registry as described with reference to Figure 7.

[0130] It should be noted that the method described herein describes a possible implementation, and that the operations and the steps may be rearranged or otherwise modified and that other implementations are possible.

[0131] The description herein is provided to enable a person having ordinary skill in the art to make or use the disclosure. Various modifications to the disclosure will be apparent to a person having ordinary skill in the art, and the generic principles defined herein may be applied to other variations without departing from the scope of the disclosure. Thus, the disclosure is not limited to the examples and designs described herein but is to be accorded the broadest scope consistent with the principles and novel features disclosed herein.