3GPP TSG SA Meeting #102 SP-231678\_rev1

11-15 December 2023, Edinburgh, UK (revision of SP-231371)

**Source: Qualcomm (Moderator)**

**Title: New SID on Core Network Enhanced Support for Artificial Intelligence (AI)/Machine Learning (ML)**

**Document for: Approval**

**Agenda Item: 6.4.2**

3GPP™ Work Item Description

Information on Work Items can be found at <http://www.3gpp.org/Work-Items>   
See also the [3GPP Working Procedures](http://www.3gpp.org/specifications-groups/working-procedures), article 39 and the TSG Working Methods in [3GPP TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm)

Title: Study on Core Network Enhanced Support for Artificial Intelligence (AI)/Machine Learning (ML)

Acronym: FS\_AIML\_CN

Unique identifier :

Potential target Release: Rel-19

# 1 Impacts

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Affects: | UICC apps | ME | AN | CN | Others (specify) |
| Yes |  |  |  | X |  |
| No | X (NOTE 1) | X (NOTE 1) | X (NOTE 1) |  |  |
| Don’t know |  |  |  |  |  |

NOTE 1: Until there is clear information on what is required from AN, it is assumed that there are no impacts on the AN/UE.

# 2 Classification of the Work Item and linked work items

## 2.1 Primary classification

### This work item is a …

|  |  |
| --- | --- |
| X | Study |
|  | Normative – Stage 1 |
|  | Normative – Stage 2 |
|  | Normative – Stage 3 |
|  | Normative – Other\* |

**\* Other = e.g. testing**

## 2.2 Parent Work Item

For a brand-new topic, use “N/A” in the table below. Otherwise indicate the parent Work Item.

|  |  |  |  |
| --- | --- | --- | --- |
| Parent Work / Study Items | | | |
| Acronym | Working Group | Unique ID | Title (as in 3GPP Work Plan) |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

### 2.3 Other related Work Items and dependencies

|  |  |  |
| --- | --- | --- |
| Other related Work /Study Items (if any) | | |
| Unique ID | Title | Nature of relationship |
| 940084 | Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR Air Interface | Related study for RAN intelligence |
| 940073 | Study on Enablers for Network Automation for 5G – phase 3 | Related study for 5GC intelligence |
| 980019 | (Stage 2 for AIML) System Support for AI/ML-based Services | Related study for AIML |
| 941010 | Artificial Intelligence (AI)/Machine Learning (ML) for NG-RAN | Related study for AIML NGRAN aspects. |

# 3 Justification

The convergence of communication network and Artificial Intelligence (AI) technology is progressing based on work done in previous releases on AI-enabled network architecture and leveraging AI/Machine Learning (ML) to enable 5GC and air interface intelligence in terms of data collection, ML model training, analytics inference, and closed-loop procedures by consuming data analytics, etc. NWDAF mechanisms and enhancements have been defined in Rel. 16, Rel. 17, and Rel. 18. Moreover, Rel. 18 AIMLsys has introduced 5GC assistance capabilities to support AI/ML operations in the application layer specified in TS 22.261.

Based on the above, for Rel. 19 it is proposed to expand the scope of network AI services to leverage AI/ML technologies to enable 5GC and Air interface Intelligence by providing network automation and improving the efficiency of the 5G network architecture.

Two main components are considered for Rel. 19:

* AI/ML alignment and convergence for Air interface and 5G Core network
* Architecture enhancement to support 5G Core intelligence.

This study item builds on Rel-18 work to support enhancements to 5GC intelligence, alignment and convergence between SA2 and RAN WG for UE data collection framework, and ML model sharing. Based on the work in previous release and the work in RAN, there is a need to study how to leverage AI/ML technologies to enable 5GC and air interface intelligence by providing network automation and improving 5G system efficiency. This may include handling of signaling storms, 5GC NF operations (i.e. policy control and QoS) assisted by NWDAF, and support for AI/ML for air interface.

AI/ML for NR air interface is being discussed in RAN1 and RAN2 in Rel-18. RAN2 is leading the discussion about data collection from UE for ML model training and ML model transfer/delivery to UE. Several solutions have been discussed in RAN2, and work in SA2 is required to analyze and discuss aspects that go beyond RAN scope. The required work includes studying how the 5GS architecture can be enhanced to enable AI/ML Radio, including the support of a common 5GS AI/ML framework to enable AI/ML Radio, whether and how to enhance the UE data collection framework to meet requirements for AI/ML model training for RAN, whether and how to support model transfer/delivery to the UE.

In addition, RAN3 has specified support for AI/ML at the NG-RAN over the course of past releases to address requirements of use cases focused on mobility optimization, energy savings and load balancing. SA2 also specified support for some NWDAF analytics that may be related to those use cases, but in both cases no cross-domain exposure capabilities have been specified. Therefore, some level of coordination between 5GC and NG-RAN may be desired to avoid redundant or conflicting AI/ML-assisted decisions in the network.

There are widespread efforts and support for AI/ML across various 3GPP WGs, and thus work is required to study possible architectural and functional extensions for cross-domain AI/ML interworking and coordination (e.g., UE, RAN, Core, applications, etc.) to address the overall AI/ML framework. This includes studying whether and how Vertical Federated Learning (VFL) may support AI training and inference for verifying and/or predicting e.g. the application QoE performance, network energy saving, and mobility optimization, and to support the AI/ML related data or model sharing.

# 4 Objective

The aim of this study work is to investigate and identify potential architecture and system level enhancements to support AI/ML enhancements.

Specifically, the objectives include:

- WT#1: AI/ML cross-domain coordination aspects

Study enhancements to support AI enabled RAN based on conclusions of the RAN study in 3GPP TR 38.843. The WT will discuss whether and how to support the cross domain (i.e. UE, RAN, 5GC, OAM and AF) collaborative AI/ML mechanisms for the aspects described by the work tasks below.

- WT1.1 –Study whether and how to support UE data collection to meet requirements for RAN AI support for air interface operation (for RAN) considering the proposal 2 (i.e. “UE collects training data and transfers it to Core Network. ~~Core Network transfers the training data to the OTT server.~~”) for UE-side model training documented in 3GPP TR 38.843 clause 7.2.1.3.2 if the relevant RAN WG concludes this option. This includes identifying what benefit can be achieved from enhanced UE data collection in 5G System, and the potential impacts on the 5G framework, including potential enhancements to policy control. The WT will also discuss the possible data leakage from the operator’s domain which should be avoided. The Core Network does not send the training data to the OTT server.

- WT1.2 –Study whether (and how) to support model transfer/delivery to the UE according to RAN1/RAN2 considerations documented in 3GPP TR 38.843 clause 7.1.2.4 , including potential enhancements to policy control if the relevant RAN WG conclude the option that 5GC is impacted. Whether and what entities or functions transfer the AI/ML model or information to the UE will be studied as part of the work. This WT will also discuss the possible data leakage from the operator’s domain which should be avoided.

**-** WT1.3: Void

- WT1.4: Void.

- WT1.5: Study whether and how to consider enhancements to LCS to support AI/ML based Positioning considering the conclusions in 3GPP TR 38.843.

NOTE A: The work will not modify the architectural principle that a service-based architecture only applies for 5GC.

NOTE B: SA2 can start WT 1.1, 1.2, 1.3 and 1.5 only if there are explicit LS response(s) from the involved RAN WG(s) by XXX 2024 that have made requirements within scope of aforementioned WTs for concluding [starting] SA2 work.

NOTE C: Further alignment with SA5 for the AI/ML Functional framework and charging may be required.

NOTE D: security aspects are in the scope of SA3, however architectural aspects related to security enhancements will be discussed in this WT.

NOTE E: The model management will follow the framework as defined by RAN.

* WT2: Study whether and what potential enhancements are needed to enable 5G system to assist in collaborative AI/ML operation involving 5GC/NWDAF and/or AF for “Vertical Federated Learning (VFL)”. The work will be based only on and limited to the scope of justified use cases.

NOTE F: RAN and UE aspects are out of scope. Solutions based on interactions between the application client and 5GS are out of scope. The necessary communication between AF and UE application client to support the collaborative AI/ML operation is understood as no normative procedure impact. Horizontal FL procedure defined in R18 should be taken into account and reused whenever possible.

NOTE G: coordination with SA6 is required.

* WT3: Study enhancements to support NWDAF-assisted policy control and address network abnormal behaviour
* WT3.1 – Study whether and what additionally needs to be supported in order to enhance 5GC NF operations (i.e. policy control and QoS) assisted by NWDAF. The work will firstly identify the specific use cases to be considered, in order to identify the appropriate scope. The work will analyse the result impacts on NWDAF (e.g. the need to understand specific NF functionality), and the compatibility of new solutions wrt existing analytics, in order to determine the need and benefits of new solutions.
* WT3.2 – Study prediction, detection, prevention, and mitigation of network abnormal behaviours i.e. signalling storm with the assistance of NWDAF.

NOTE H: The study will focus primarily on existing enforcement mechanisms when available and identify new ones when no existing ones can be used.

NOTE I: The study will consider the study/work done by SA WG5 and CT WG4 in this regard already and collaborate with SA WG5/CT WG4 regarding the handling of abnormal network behaviours.

## TU estimates and dependencies

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Work Task ID** | **TU Estimate**  **(Study)** | **TU Estimate**  **(Normative)** | **RAN Dependency**  **(Yes/No/Maybe)** | **Inter Work Tasks Dependency** |
| WT#1 | 4 (see Note J) | 4 (see Note J) | Yes |  |
| WT#1.1 | 1 | 1 |  |  |
| WT#1.2 | 1 | 1 |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
| WT#1.5 | 1 | 1 |  |  |
| WT#2 | 1 | 1 | No |  |
| WT#3 | 2 | 2 | No |  |
| WT#3.1 | 1 | 1 |  |  |
| WT#3.2 | 1 | 1 |  |  |

NOTE J: SA2 can study WT1 only if there is explicit LS response(s) from the involved RAN WG(s) that have made requirements within scope of WT1 for starting SA2 work. For WT1, the specific number of TUs can be determined after finalizing the WT1 description required by the LS response(s) from RAN WG(s). Though the overall TUs estimated for WT1 are 8 as an absolute upper limit, but it would be reduced based on LS response(s) from the relevant RAN WG(s) including the explicit requirements for SA2 to start the work. The respective TU estimation for each sub-WT under WT1 can be determined after finalizing the WT description.

**Total TU estimates for the study phase: 7**

**Total TU estimates for the normative phase: 7**

**Total TU estimates: 14 (see NOTE J)**

# 5 Expected Output and Time scale

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| New specifications {One line per specification. Create/delete lines as needed} | | | | | |
| Type | TS/TR number | Title | For info  at TSG# | For approval at TSG# | Rapporteur |
| Internal TR | 23.xxx | Study on Core Network Enhanced Support for Artificial Intelligence (AI)/Machine Learning (ML) | TSG#103  June 2024 | TSG#104  June 2024 | TBD |

|  |  |  |  |
| --- | --- | --- | --- |
| Impacted existing TS/TR {One line per specification. Create/delete lines as needed} | | | |
| TS/TR No. | Description of change | Target completion plenary# | Remarks |
|  |  |  |  |
|  |  |  |  |

# 6 Work item Rapporteur(s)

TBD

# 7 Work item leadership

S2

# 8 Aspects that involve other WGs

The following aspects involving other WGs may arise related to this SID:

- Support for security and privacy aspects on data collection and model transfer/delivery to UE should be discussed in SA3.

- RAN aspects

- SA4 for EVEX and AI4Media

- SA5 for MDAS, AI/ML and charging

- CT4 for signalling storm aspects

- SA6 for application client related aspects

# 9 Supporting Individual Members

|  |
| --- |
| Supporting IM name |
| China Mobile |
| ETRI |
| FirstNet |
| Interdigital |
| Lenovo |
| KDDI |
| KPN |
| OPPO |
| Rakuten Mobile |
| SK Telecom |
| Vivo |
| Deutsche Telekom |
| Telefonica |
| Verizon |
| BT |
| CATT |
| China Telecom |
| China Unicom |