**SA WG2 Meeting #161S2-2401968**

**February 26 – March 1, 2024, Athens, Greece**

**Source: Samsung, OPPO, ZTE, Futurewei, Lenovo, ETRI, NTT Docomo, SK Telecom**

**Title: WT2: New Use Case for Vertical Federated Learning**

**Document for: Approval**

**Agenda Item: 19.15**

**Work Item / Release:** **FS\_AIML\_CN / Rel-19**

*Abstract of the contribution: This pCR proposes a new use case for KI#2 on NWDAF Support for Sample and Feature Alignment in Analytics.*

# 1 Discussion

This paper proposes a use case for KI#2: 5GC Support for Vertical Federated Learning.

# 2. Proposal

It is proposed to adopt the following text in TR 23.700-84.

\*\*\* Start of 1st change (all new text) \*\*\*

## 5.1.X Use Case #X: Support for VFL in 5GC

It is well known in the AI/ML literature that VFL is a federated learning setting where multiple parties perform training on data sets that share the same sample space but differ in feature space. Because of this, an alignment in sample and feature spaces among participating entities is usually required before applying VFL. VFL further allows to perform joint training without exposing raw data, with each entity owning its own model but not needing the same model architectures. This is a way in which VFL differs from HFL. TS 23.288 [X] provides NWDAF specification support for HFL but no VFL support is available, and the existing procedures defined for HFL may be enhanced to support VFL, as mentioned in cl. 5.2.4.

This use case proposes to support VFL in 5GC for analytics derivation leveraging sample and optionally (if needed) feature alignment between the entities participating in VFL, and where the main entity facilitating the VFL operation is NWDAF and other entities may be other NWDAF instances. In a multi-vendor scenario, this would allow participating NWDAF instances to collaborate in VFL without the need for model parameter exchange.

NOTE: This use case does not propose specifying new multi-vendor capabilities for NWDAF. But rather solve the Interoperability of Intermediate results between vendor instances.

The motivation to introduce VFL in 5GC is as follows:

- Enables to perform distributed inference: VFL enables to perform distributed inference, as such the ML model does not need to reside in one place as in HLF, and given that inference is distributed, it is possible that the deployment enables the access to the input data close to the source of data as such it further reduces signalling and computing time, the ML model does not need to reside in one place as for HFL.

- Reduces the need for offline interoperability testing; In HFL the expectation is that interoperability testing is required to check that NWDAF from different vendors can interoperate, i.e., the Vendor-id can only work if there is a prior interop testing between vendors. In VFL the need for offline interoperability will be studied, it needs to be resolved, however there is no need to test that models from different vendors can interoperate, but the interoperability of intermediate results provided in each VFL interaction may be enough.

- A more flexible technique: In VFL vendor specific local models and features can be deployed in each participant, so that it is possible that each participant selects the local model to be used unless constraints apply, as such vendor or operator specific local models and features, including not standardized features, are simpler to implement comparing with HFL.

In one PLMN where multiple NWDAFs are deployed, each NWDAF instance may perform data collection according to their available data sources. Depending on the Analytics ID and the deployment scenario however, the different NWDAF instances may share the same sample space or train on different sample spaces. VFL would be beneficial on the former case. Furthermore, in VFL each NWDAF instance does not collect the same input data for the same Analytics ID, thus their feature spaces may range from full to little or no overlap.

\*\*\* End of 1st change \*\*\*