**3GPP TSG-SA3 Meeting #115 AdHoc-e *S3-241312-r2***

**Electronic meeting, online, 15 - 19 April 2024**

**Source: Huawei, HiSilicon, vivo**

**Title: New KI on** **UE ID security and privacy of VFL between VFL members.**

**Document for: Approval**

**Agenda Item: 5.13**

# 1 Decision/action requested

***approve this new KI for inclusion in TR 33.784.***

# 2 References

[1] 3GPP TR 23.700-84, ' Study on Core Network Enhanced Support for Artificial Intelligence (AI)/Machine Learning (ML) '.

[2] 3GPP TS 23.288, 'Architecture enhancements for 5G System (5GS) to support network data analytics services', (Release 18).

[3] "IEEE Guide for Architectural Framework and Application of Federated Machine Learning," in IEEE Std 3652.1-2020.

# 3 Rationale

As description Clause 5.1.5, in SID TR AIML [1], there is a Use Case for NWDAF support for observed service experience analytics based on VFL. It provides the real user feedback to the network so that the network could self-optimize and offer customized services according to the true user needs.

In clause 6.4.4 of TS 23.288 [2], as description of Analytics ID for observed service experience, the consumer needs to request the Analytics ID "Service Experience" for a UE identified by a SUPI or a group of UEs identified by a list of Internal Group-Ids in the procedure. In this case, the UE ID (e.g, SUPI, GPSI, AF specific UE identifier) may need to be sharing within the VFL participants (e.g., NWDAF(s), AF(s)). However, The AF specific UE identifier is ensured to be unique across different AFs as defined in TS 23.003, which means the different AF have different UE IDs for the specific UE. For example, UE1 is both user of AF1 and AF2. The AF UE ID of UE1 on AF1 is different from the AF UE ID of UE1 on AF2.

As description in IEEE Guide for FL [3], Sample alignment module, A sample alignment module is mainly used for vertical federated machine learning. The module identifies the overlapped samples of different data sources and does not disclose sample feature information.

How to maintain UE ID security and privacy when NWDAF(s) and AF(s) doing VFL.

# 4 Detailed proposals

SA3 is kindly requested to agree the pCR.

\*\*\* BEGINNING OF 1st CHANGES \*\*\*

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[xx] 3GPP TR 23.700-84, ' Study on Core Network Enhanced Support for Artificial Intelligence (AI)/Machine Learning (ML) '.

[yy] 3GPP TS 23.501, "System architecture for the 5G System (5GS) ".

[zz] 3GPP TS 23.003: "Numbering, Addressing and Identification".

[aa] "IEEE Guide for Architectural Framework and Application of Federated Machine Learning," in IEEE Std 3652.1-2020.

[bb] 3GPP TS 23.288, "Architecture enhancements for 5G System (5GS) to support network data analytics services".

\*\*\* BEGINNING OF 2st CHANGES \*\*\*

# 5 Key issues

Editor’s Note: This clause contains all the key issues identified during the study.

## 5.X Key Issue #X: Privacy of VFL between VFL members.

### 5.X.1 Description

Vertical federated learning (VFL) allows the cooperation of multiple NWDAF(s) and/or AF(s) to cooperate to train models locally where no raw data need to be exchanged.

As description in IEEE Guide for FL [aa], Sample alignment module, a sample alignment module is mainly used for vertical federated machine learning. The module identifies the overlapped samples of different data sources and does not disclose sample feature information.

The sample alignment procedure may involve the exchange of information (e.g, UE ID) which is sensitive and could potentially comprise the privacy of UEs.

The VFL data (e.g, intermediate result) which could be used by comporised members to do Label or feature inference attack. Consequently, these data should be protected and authorized.

#### 5.X.2 Security threats

The UE ID privacy may be leaked between VFL participants from different domains when doing VFL, the AF may obtain UE information (UE ID) supported by another AF.

#### 5.X.3 Potential security requirements

5GS shall support privacy protection on sample alignment procedure.

Editor’s Note: The terms can be updated based on SA2’s progress.