r4

Variant of RAN93e-R18Prep-12 AI/ML email discussion Sept 2021 Version 0.0.4

RAN

3GPP TSG RAN#93e **RP-211662**

Electronic Meeting, Sept 13 – 17, 2021

Agenda item: 9.0

Source: Moderator (Qualcomm)

Title: Moderator's summary for discussion [RAN93e-R18Prep-12] AI/ML

Document for: Report

1 Initial Round

1.1 General high-level views

Companies views:

Feedback Form 1: General high-level views

1 – vivo Mobile Communication Co.

We fully support this study as shown in our previous papers.

To make the study more future-proof, we believe an initial phase for use case study is needed. This consideration is based on the following two aspects:

1. The study would satisfy long term commercial needs. Without a study on whole picture of potential use cases in Rel-18 and beyond, the studied aspects would be limited to specific areas. This would potentially lead to different solutions in different phases throughout the lifetime of 5G standardization, which would not be helpful for potential deployment in future.

2. The use case that can be discussed before December would be high level. Within each specific area, the use case in people's mind would be different, e.g., some companies believe CSI is about CSI reporting overhead compression, others believe CSI is about how to enable more CSI report to facilitate training. The expected study and specification work would be completely different for these different understandings. Without a phase to clarify what use cases are expected, the TU allocation and study arrangement would be hard to manage.

2 – Futurewei

- Futurewei fully supports setting up a study item in Rel-18 focusing on AI/ML for PHY.
- To understand AI/ML for PHY, we suggest the following:

o Identify high-level principles for AI/ML-based operations in PHY, leveraging the RAN intelligence framework under Rel-17 RAN3 SI, FS_NR_ENDC_data_collect.

o Identify few candidate use cases that would benefit from leveraging AI/ML technology based on companies' studies.

o Study common methodologies for performance evaluation.

3 – InterDigital

We support the study. Since AIML is a new area for 3GPP, we think that the initial study phase should focus on evaluation methodology including KPIs, data sets for training and verification, assumptions on UE capability, and characterization of AI model performance.

4 – BT plc

We welcome a study into potential practical enhancements, feasibility, and requirements on standardisation from effective interoperability of AI/ML-related enhancements. Such study should start with a focus on the challenges AI/ML could be solving effectively, both in shorter-term and longer-term.

5 – Apple Computer Trading Co. Ltd

We fully support this study. We would like to clarify the leading working group.

- For AI/ML for air interface, this is RAN1 led SI. Related signaling/procedures involving RAN2/RAN3 should be studied as well.

- For AI/ML for RAN, this is RAN3 led WI. Related signaling/procedures involving RAN2 are part of the WI.

6 – Qualcomm CDMA Technologies

We are supportive of looking into ways to apply AI/ML techniques into wireless communication problems involving air-interface and going beyond what current implementation based approaches can achieve, which, do not rely on any inter-node exchange of information. For these approaches, it would be important to identify the potential performance gains over classical procedures/implementations along with a good understanding of the required air-interface impact, as well as, the corresponding performance and complexity.

We see the possibility of looking into various levels of inter-node collaboration and identifying their respective performance, complexity and required air-interface impact.

A few examples of those various levels of collaboration follow:

0) No collaboration framework: AI/ML algorithms purely implementation based and not requiring air-interface changes/extensions.

1) Inter-node assistance to improve the respective nodes AI/ML algorithms (not currently possible with current NR framework).

2) Joint ML operation between UEs and network identifying what the air-interface impact would be.

3) Hybrid approach with pieces of 0), 1) and 2) above.

Equally important, we believe that the evaluation of AI/ML techniques should not purely based on synthesized channel models, interference, noise, etc. Indeed, we believe that the introduction of real world traces for the corresponding performance evaluations is important and should complement/augment any analysis carried out based on model-based simulations.

7 – SHARP Corporation

We are totally positive to study AI/ML for air interface in Rel-18. Along with the development and maturing of AI/ML techniques, there are tons of research works demonstrating the benefits of applying AI/ML in

wireless communications, at various aspects. Meanwhile, compared to legacy systems, 5G NR provides much more flexibility for air interface, which makes it feasible and natural to apply AI/ML. We believe study item AI/ML for air interface in Rel-18 would be a good starting point, where AI/ML application can be enabled with minor specification change on the existing 5G NR standardizations. For long term considerations, applying AI/ML for air interface should be an essential factor when designing the next generation system, e.g., 6G.

8 – Xiaomi Communications

Xiaomi

- The high level principle defined for project of "FS*NR*ENDC*data*collect" in RAN3 can set as a starting point
- Setting a short-term SI phase to reach contain consensus in ML/AI background, e.g., algorithm, training manners would be helpful
- Basic principle should be discussed for the data set construction
- Impact on the devices complexity, power saving should be taken into account as well

9 – ZTE Corporation

In general, we think AI should be core functionality for 5G advanced, including both AI for air inter face and AI for NG-RAN.

- Follow-up normative work should be setup in Rel-18 considering the popular AI use cases identified in RAN3 focusing higher layer impacts.

- General benefit and performance gain can be observed in primary research from multiple sources. Study on AI applications in physical layer can be considered to start the AI standardization work on PHY for 5G Advanced.

10 – Intel Technology India Pvt Ltd

- Important to define a focused air-interface study commensurate with TU budget, otherwise WG discussions can be inefficient.
- Consider methodologies for performance requirements and testing [RAN4] for AI/ML. Some discussion should start in RAN4 within the air-interface SI term.
- High-level use-case selection in RAN plenary is preferred for air-interface (it is complex and time-consuming to do a similar exercise in RAN1 as use-cases are coupled with EVM and specifications)
- For air-interface study focus on supervised and centralized learning evaluations

11 – Deutsche Telekom AG

We support several activities related to introduction of AI/ML approaches in RAN within Rel-18.

The first one is a WI based on the outcome of the Rel-17 SI performed in RAN3. This WI should consider the functional framework for RAN intelligence within NG-RAN as described in TR 38.817 and the use cases covered in that TR.

The impact of further use cases can be studied in a new SI which, in addition, should also study and identify the possibilities for base such a framework on Service Based Architecture (SBA) principles, which are already used in 5GC and OAM, and to enable a soft migration into a fully cloudified environment for the entire Data Analytics and ML/AI incorporating NWDAF and MDAS approaches of SA2 and SA5.

Furthermore, a SI should be performed AI/ML approaches for the air interface. Due to the complex environment we don't see the need for a follow-up WI phase within Rel-18.

12 – Nokia Corporation

AI/ML for Air interface:

Diverse AI/ML proposals for the air interface were made in the workshop. Most of the proposals require fundamental studies to ensure that system benefits and gains are obtained, and predictable device behaviour and performance can be guaranteed. Hence, future work in this area will require a relatively long SID, with care to engage all relevant WGs (RAN1/2/4) early on.

13 – NTT DOCOMO INC.

We support the study of AI/ML for air interface in Rel-18. Since this is the first step to introduce AL/ML for air interface, use cases/enhancements with huge specification impacts should be avoided. Also, we prefer down-selecting use cases to reduce workload for study item. As an evaluation methodology depends on AI/ML use case, it might be necessary to set up as many evaluation methodologies as AI/ML use cases to study. We believe having moderate number of use cases leads to success in AI/ML for air interface in Rel-18.

14 – CATT

AI is considered to be an important component of 5G-Advanced from both NG-RAN and air interface perspective. For AI application to air interface, studies on KPI, evaluation methodology and construction of data set shall be prioritized. Studies laid out in Rel-18 could build foundations for potential standardization in future releases.As to AI for NG-RAN,work could be continued based on the outcome of RAN3 leading Rel-17 SI.

15 – CAICT

We support the study of AI/ML for air interface and specification work of AI/ML for NG-RAN in R18.

The agreements achieved in NG-RAN study could be used for the starting point for AI/ML for air interface and detailed AI/ML model should be left for implementation. In order to support AI/ML for both air interface and higher layer, unified network structure should be considered.

AI/ML for air interface study requires performance evaluation. An initial phase for use case study is helpful.

16 – China Mobile Com. Corporation

AI for NG-RAN:

A Rel-18 WI following on the Rel-17 SI should be carried out. It could follow the conclusions on AI framework and the prioritized use cases in the Rel-17 TR. RAN3 will be leading group with RAN2 involvement as secondary WG. Further study on other use cases can be performed in parallel.

AI for air interface:

We support the study of AI/ML in Rel-18 and think evaluation methodology is the most important issue. The initial study phase should make clear the evaluation methodology and focus on a few use cases.

There will be potentially a study focusing on AI for air interface (RAN1 leading with RAN2/RAN3 involvement) and also a further study on other AI use cases for NG-RAN (RAN3-leading with RAN2 involvement), so how to organize the Rel-18 study on AI enabled NG-RAN and air interface should be further discussed, e.g., whether to include them into one study item?

17 – Fraunhofer IIS

We support this study. Both AI for Air Interface and RAN-NG has a lot of potential for improving and/or enabling usage of the technology.

18 – TELECOM ITALIA S.p.A.

We share the view expressed by Deutsche Telekom. An initial Work Item to ensure that data can be collected followed by a long term study for a specification phase in Rel 19 is our preferred way forward.

To ensure data collection (Work Item phase), features like MDT should be made mandatory in the UE and proper interfaces / data formats should be specified from the network side.

19 – Spreadtrum Communications

We support to kick off the study work on AI/ML enabled RAN including air interface, and one WID on AI/ML applicable in high layer based on R17 SID. Regarding how to carry out the work, one way is to set up the study in RAN level, and package High layer and air interface; another way is to set up the study in respective WG, e.g., one SI in WG1, and one WI in WG2/WG3. The former aims to provide common basis, e.g., AI/ML terminology, dataset construction, AI/ML model, and so on. The latter can focus specific use cases to trigger discussion. We are ok for either way.

20 – Guangdong OPPO Mobile Telecom.

We are supportive of having the AI/ML study item in Rel18. The objective of this study item should at least include:

- Study evaluation methodology for AI/ML model-based solutions, incl. training data set, evaluation metrics, generalization characteristics, complexity impacts.
- Study the AI/ML model-based enhancements in some use cases (e.g. up to 2), and evaluate the performance gain over the traditional algorithm.
- Study the potential enhancements for supporting AI/ML model transfer (AMMT) and AI/ML services in 5G system.

21 – NEC Corporation

We support both AI/ML for Air Interface and AI/ML for NG-RAN directions. Also, in Release 18 integration/alignment/joint operation of OAM AI/ML, 5GC AI/ML, NG-RAN AI/ML, and Air Interface AI/ML could be considered, as well as, OAM management aspects of NG-RAN AI/ML and Air Interface AI/ML.

22 – Samsung R&D Institute UK

- We expect this new SI evaluate practical gain and feasibility of applying AI in 5G NR PHY:
  - Study appropriate evaluation methodology (training based)
  - Study appropriate framework
  - Identify key use cases, e.g., CSI feedback, Beam management, non-linear handling
  - Identify the necessary specification support to facilitating AI in PHY, even though AI itself is black box and implementation specific
    - E.g. how to collect data-sources for online training, which may impact **signal/channel** design, measurement, feedback/report etc.

23 – Fujitsu Limited

•We support to set up AI for air interface as a RAN1-led study item in Rel-18.

•High level principles defined in "FS*NR*ENDC*data*collect" of RAN3 should be taken as the reference for the study of AI for air interface. Further discussions on other principles related to feasibilities of using AI with real time constraints from the focused use cases.

•Besides, it is suggested that companies to have a common understanding on AI terminology, functionalities, framework, and evaluation methodology during the initial study phase.

24 – KDDI Corporation

a) Air interface

We are supportive the study of AI/ML for air interface in Rel-18. From air interface point of view, this study of AI/ML is the first step. Therefore, we think that it is an important to firstly focus on specific use cases as well as RAN3 has already done in Rel-17 AI/ML SI.

b) NG-RAN

We have same views with CMCC and also think that progress and outcomes from RAN3 in Rel-17 can be considered as a baseline. Considering work load and Rel-17 SI on NG-RAN has been started, it would be better that air interface of AI/ML is prior to NG-RAN in Rel-18.

25 – Lenovo (Beijing) Ltd

We support to have the studies of AI/ML-assisted enhancement for air interface and evolution of AI/ML for NG-RAN.

To have a clear study scope with solid conclusion, we suggest carrying out the studies phase-by-phase, and in the initial phase, it is necessary to have a common understanding on the evaluation methodology for the AI/ML-assisted approaches, which can be basic for the long-term evaluation in following phases.

26 – HUAWEI TECHNOLOGIES Co. Ltd.

In Rel-17 RAN3 study item for data collection, some high-level principles have been discussed and agreed for AI/ML. Our view is that these high-level principles shall also be applied for the potential study of Rel-18 PHY AI/ML. Such high-level principles include:

1□Detailed AI/ML algorithms and models are left for implementation;

2□User data privacy needs to be preserved;

3□Support of AL/ML shall reuse the existing RAN architecture and new interfaces shall not be introduced.

Overall, our view is that PHY AI/ML is a promising new area that deserves a thorough study first. Thus, we think it is reasonable to only target a study item on PHY AL/ML in Rel-18.

27 – VODAFONE Group Plc

We are generally supportive of the views from DT and Telecom Italia.

28 – MediaTek Inc.

We generally support the study on AI/ML integration with air interface and RAN. We also understand such integration is a long term process with wide scope and therefore would like to start the work in all related WGs but with selective use cases.

29 – Sony Corporation

AI/ML for air interface could improve efficiency of wireless systems. So, we think that it will be worth studying AI/ML for air interface in terms of use case(s), evaluation metrics, how and which entity implements the AI/ML model, and others in a new study item of AI/ML for air interface. Below are possible objective of the new study item.

- Identify potential use cases for AI / ML
- Identify use cases with a standards impact
- Identify metrics / KPIs for evaluating AI / ML
- Compare complexity (and other metrics, such as UE power consumption) of an AI/ML based solution to a ”classic” approach
- Determine what is required to allow for AI / ML training
- Determine what is required to support AI / ML inference
- Take some example use cases (for example CSI enhancement) and look into that use case in more detail

30 – Ericsson LM

Ericsson is supportive of a Rel-18 AI/ML for PHY study item. We propose the following objectives of such Rel.18 study item:

- Objective 1: Study, at a high level, methodologies to enable specification of AI/ML-enhancements for physical-layer use cases.
- Objective 2: Complete a detailed (example) study for the specific application of AI/ML-enhanced CSI reporting to learn how AI/ML impacts on PHY specifications.
- Objective 3: Document the general learnings in a technical report (TR), to create a template for future AI on PHY use cases in 3GPP.

31 – MediaTek Inc.

We generally support the study of AI/ML integration with air interface and RAN. We also understand the vast scope and long term migration of such integration, therefore, would like to start work in all related WGs but with selective use cases first.

32 – Ericsson LM

Ericsson is supportive of a Rel-18 AI/ML for PHY study item. We propose the following objectives of such Rel.18 study item:

- Objective 1: Study, at a high level, methodologies to enable specification of AI/ML-enhancements for physical-layer use cases.
- Objective 2: Complete a detailed (example) study for the specific application of AI/ML-enhanced CSI reporting to learn how AI/ML impacts on PHY specificaitons.
- Objective 3: Document the general learnings in a technical report (TR), to create a template for future AI on PHY use cases in 3GPP.

33 – AT&T

Areas for potential enhancement include application awareness regarding the specific mobility conditions of XR users and L1/L2 inter-cell mobility enhancements. In addition, considerations of potential DL/UL imbalance during mobility scenarios could be of particular interest for XR users since both directions need to be optimized considering the overall end-to-end user experience.

34 – Orange

We are supportive of a WI for AI/ML related to SON / NG RAN in Rel18, and a SI for AI/ML related to air interface optimisation.

35 – Verizon UK Ltd

We support views of Deutsche Telekom and Telecom Italia. We see AI/ML is one of the key technologies in Rel-18 and envision the following activities:

1) A WI for AI/ML in NG RAN based on current RAN3 SI (basic AI/ML framework and limited set of use cases)

2) A SI focussed on air interface use cases

3) A SI for AI/ML for NG-RAN that considers further framework enhancements and use cases

All relevant WGs RAN1/2/3/4 should be involved as appropriate in these activities

1.2 AI/ML for Air Interface: Use cases to focus

Companies views:

Feedback Form 2: AI/ML for Air Interface: Use cases to focus

1 – vivo Mobile Communication Co.

At least we have the following areas in mind:

- Positioning performance improvement under various challenging conditions;
- CSI feedback with lower overhead to achieve higher system efficiency;
- RS overhead reduction and channel estimation performance improvement, including DMRS/CSI-RS/SRS;
- Intra-cell/Inter-cell beam management latency and accuracy improvement
- Mobility performance enhancement, including power consumption reduction and latency reduction.

In each of the specific areas, we believe there are different use cases including the following different kind of use cases:

- Cat1: AI/ML related training and inference are all conducted at one side of network or UE and is transparent to the other side;
- Cat2: AI/ML related training and inference are conducted at one side of network or UE, but requires additional signaling or procedure enhancements between two sides, potentially with existing signaling framework. Additional information is not directly related to training and inference, e.g., capability,

new patterns etc.; Training or inference is done by implementation and not explicitly seen in the specification

- Cat3: AI/ML related inference is conducted at one side of network or UE, with assisted training information exchanged between two sides; E.g., neural network models provided as a reference for UE or for aggregation
- Cat4: AI/ML related inference are conducted together at both sides of network and UE training maybe conducted at one side or both. Information related to inference need to be exchanged between both sides;

For example, for positioning related areas, we see different use cases including

- Cat1 use case for positioning: AI/ML based OTDOA estimate at the UE with training and inference transparent to the NW;
- Cat2 use case for positioning: Additional gNB antenna information/calibration information can be provided to UE to improve positioning estimate accuracy based on AL/ML;
- Cat3 use case for positioning: Neural network trained/aggregated at network for the environment are distributed to UE to expedite the training at the UE;
- Cat 4 use case for positioning: Inference at UE side to extract features and reported to network that would further use by NW for inference.

For most of the listed areas, we see above different kinds of use cases in each of the area.

2 – Futurewei

· Futurewei's view is to study the benefit of AI/ML for PHY via the following use cases:

a) AI/ML-based CSI feedback compression and reconstruction to reduce CSI feedback overhead while achieving high accuracy in CSI reconstruction.

b) AI/ML-based channel prediction, e.g., predicting subsequent channels by exploiting the temporal correlation in channel history.

c) AI/ML-based beam selection, e.g., using the received signals at Rx, adaptively find the beamforming vector that maximize the channel capacity.

3 – InterDigital

In terms of use case, we think the initial study phase should identify/down select at most one use case for which AI based solution has the most benefit and clear specification impact. It is important that we avoid specifying multiple use cases in parallel, so that learnings from AIML methodology/framework created for the initial use case can be applied for other use cases.

4 – BT plc

In our view use cases to focus depend on (1) problems AI/ML will be solving, and (2) how standardisation can support interoperability of techniques benefitting from AI/ML methods to solve challenges in (1).

With respect to (1) above (problems worth solving with AI) we propose to identify and evaluate potential enhancements in both short/medium and forward-looking market readiness:

1a. Enhancing current system architectures, e.g. CSI, RS aspects;

1b. Feasibility/benefit of new system architectures, e.g. new PHY/fundamental operating principles that have potential going beyond component-specific enhancements.

With respect to (2), the focus should be on approaches to enable interoperability support of efficient techniques that address both immediate and forward-looking aspects in cases 1a and 1b.

5 – Rakuten Mobile

We think, firstly we need to agree on use cases and then go into further details as requirements for each use case.

We propose to include at least following use cases

MIMO CSI Feedback Compression/Prediction, RS Overhead reduction, Predictive Mobility Management, Beam Management, Positioning enhancements, Receiver enhancements like symbol detection/channel estimation & MU-MIMO beamforming.

6 – Apple Computer Trading Co. Ltd

We prefer focus on a small set (e.g. 2-4) of prioritized use cases. Proposed high priority use cases are CSI feedback, beam management and channel estimation for RS overhead optimization.

7 – Qualcomm CDMA Technologies

From our own internal investigations and observations gathered from contributions at the 3GPP RAN Rel-18 workshop, we see beam management, cross-node CSF, RS overhead reduction (incl. channel estimation), and positioning as promising PHY layer areas to consider for AI/ML investigations. As a result, we recommend focusing on those use cases only (with the possibility for further down-selection amongst those).

8 – SHARP Corporation

We share the similar view that the study should focus on a constrained number of use cases, at the starting point of AI/ML for air interface study. The use cases should be the most important for optimizing the air interface, require less specification changes, be feasible to evaluate the performance, and be easier to demonstrate the benefits of applying AI/ML. Based on these considerations, we believe the follows could be the proper candidate use cases.

- Reference signal configuration and channel estimation
- CSI feedback compression
- Beam selection/prediction

9 – Xiaomi Communications

Generally, we should focus on uses cases achieving reasonable gain by using AI/ML-based solution with small impact on the device complexity and specification

But during the SI phase, we don't need to narrow down the use cases to focus. Companies should be encouraged to study the feasibility and the performance gain in different use cases. In our mind, the following are potential use cases

- MIMO, e.g., channel estimation, CSI processing, beam management
- Mobility management, e.g. AI utilization in higher layer procedure
- Power saving e.g., adaptive DRX management

10 – ZTE Corporation

Based on previous discussion, the following four use cases receive major interest from RAN companies.

- CSI feedback/prediction
- Beam management/mobility prediction
- Channel estimation/Reference signal overhead reduction
- Positioning

We should focus on these four use cases at least for further discussion in RAN.

11 – Intel Technology India Pvt Ltd

MIMO Use-cases

- CSI feedback focusing on PMI/RI feedback
- Beam management focusing on fast beam pair acquisition
- RS overhead reduction focusing on CSI-RS

Positioning use-cases

- LOS/NLOS classification and multi-path mitigation. Study assuming current architecture (measurements at UE/gNB side, reported to LMF)

12 – CATT

The number of use cases under study shall be limited. Involving too many use cases does not help understand how AI/ML works for physical layer. A lot of time would be spent on defining the use cases. We don't want to include too many use cases in the study.

13 – Deutsche Telekom AG

The SI should focus initially on a limited set of use cases that also allows to describe the required functional framework for AI/ML approaches across the air interface (similar what RAN3 is doing within its Rel-17 SI).

We see benefits for following use cases:

- CSI compression, channel estimation/prediction
- Beam selection/tracking
- Link adaptation (MCS selection)

14 – Nokia Corporation

During the workshop rather many diverse use cases with very different impacts on NR radio and system operations. In our view it is important that the feasibility of each AI/ML use case, application and functionality is separately assessed, and clear performance gain shown for each use case for which AI/ML is considered to be utilized. At the same time when selecting use cases it is also important to consider how we can also study means to enable feature-independent AI/ML signalling and workflow, which are used by proprietary ML implementations, so that we could better enable extensibility in a multi-vendor environment.

15 – NTT DOCOMO INC.

At least we are positive about studying following use cases:

AI CSI feedback. As collaborated ML models on both UE and NW requires large modification on the air interface from current specification, we prefer AI/ML deployments only on NW as a first step. Under this scenario, ML replaces the interpolation to dramatically improve the CSI reconstruction accuracy, while maintaining the codebook generation framework from the current air interface structure.

AI BM management. There are two approaches we have in mind: modification for optimized inputs on AI/ML (e.g., enhanced beam measurement report) and modification for utilizing inference from AI/ML (e.g., predicted beam indication). In both cases, we think use cases with AI/ML deployed only on NW should be a baseline in Rel-18.

16 – CAICT

As mentioned by vivo, we also observe that with only high-level descriptions, the understanding of different use cases is diverse. A study phase for use case identification and selection process is beneficial.

the following use cases could be study further:

- CSI feedback for overhead reduction
- Channel estimation enhancements including DMRS/CSI-RS/SRS
- Positioning enhancement
- Beam management for accuracy improvement and latency reduction
- Channel prediction
- Mobility enhancements

17 – China Mobile Com. Corporation

In our views, one or two of the potential use cases can be selected to discuss the evaluation methodology. Potential use cases can be AI/ML based CSI feedback, AI/ML based DMRS overhead reduction, AI/ML based positioning enhancement.

18 – Spreadtrum Communications

In our mind, at least the following use cases could be considered:

- AI/ML based CSI feedback with low overhead
- AI/ML based Positioning
- AI/ML based beam management and mobility

19 – Guangdong OPPO Mobile Telecom.

Considering the capacity of the study item, it is suggested to focus on limited number of use cases in Rel-18 study, e.g. up to 2 use cases. The selection of study use cases should take the following factors into account:

- Maturity of AI/ML model study in industry, academics and public 5G AI competitions – reference model is ready as a starting point;

- Maturity of training data set and evaluation methodology in industry, academics and public 5G AI competitions – data set and evaluation methodology are ready as a starting point;
- Training complexity is verified to be acceptable – Companies can afford the needed computation power;
- Cover different areas as possible when the above factors are satisfied, e.g. 1 use case for eMBB enhancement + 1 use case for vertical application.

Based on the above considerations, the primary use cases from our perspective include:

- Use case 1 (eMBB-type): AI/ML-based CSI feedback enhancements
- Use case 2 (vertical-type): AI/ML-based positioning and focus on UL based positioning and UE-assisted DL based positioning

Other use cases can also be considered, for example, AI/ML-based RS overhead reduction/channel estimation and AI/ML-based beam management, if time is allowed

20 – NEC Corporation

We are interested in AI/ML for MIMO, especially for the following use cases:

1. RS overhead reduction
2. Beam management with low overhead and latency
3. CSI feedback with low overhead
4. L1/L2 mobility

21 – Samsung R&D Institute UK

We propose to focus on at least:

- CSI feedback(including CSI compression, channel estimation)
- Beam management and tracking,
- Non-linearity handling

22 – DENSO CORPORATION

CSI and mobility predictions are potential scenarios to which we'd like to focus on at first.

23 – Fujitsu Limited

We prefer focus on several prioritized use cases, which can show obvious AI's gain and have clear standard AI impacts, listed as follows:

1. RS overhead reduction:

   - To reduce RS overhead for accurate channel estimation by utilizing AI's power in information recovery

2. CSI feedback compression:

   - To reduce feedback overhead in massive MIMO cases by using AI's feature extraction capability to compress channel state information

3. Link adaptation:

   - To overcome outdated PMI/CQI/RI Feedback in high mobility scenarios by utilizing AI's capability in time series prediction

4. Beam management

   - To have zero beam failure coverage, e.g., by utilizing AI's capabilities for decision making in uncertain and high-dimensional spaces

24 – KDDI Corporation

We agree with Futurewi. We think that the following use cases on air interface seems to be beneficial for AI/ML.

a) AI/ML-based CSI feedback

b) AI/ML-based channel prediction

c) AI/ML-based beam management and mobility

25 – Lenovo (Beijing) Ltd

We suggest selecting one or two use-case(s) in the initial phase, which should be self-contained, manageable and comparable. According to some preliminary evaluation results, we prefer to focus on the following two cases:

- CSI report with AI/ML-based compression/recovering

- Beam management with AI/ML-assisted tracking/prediction

26 – HUAWEI TECHNOLOGIES Co. Ltd.

During the past R18 workshop, many use cases have been proposed by companies, e.g., AI-assisted positioning, RS and channel estimation, CSI feedback, beam management, and so on. Our view is that only carefully selected use cases shall be included in the Rel-18 study item, as otherwise the scope and workload of the Rel-18 PHY AI/ML study item can easily explode too much. These selected use case shall be the ones where most gains from AI/ML over existing implementations are expected.

27 – VODAFONE Group Plc

We see that estimation and prediction of CSI is likely to be an application. Other use cases described by Qualcomm and Rakuten seem interesting.

28 – Ericsson LM

We see that "AI-enhanced CSI reporting" is a promising area for AI/ML based PHY that involves both gNB and UE and thus excites many different and general issues that needs to be clarified. Given limited time unit allocations and the uncertainties of this topic, we propose to focus on this one example.

If RAN1 discuss this example use case, it can provide learnings for how "AI/ML on PHY" framework might be designed in able to extend this to other use cases.

29 – Sony Corporation

We can discuss potential use cases and motivation in a new study item. And then, we can identify which use cases should be focused from outcome of new study item of AI/ML for air interface. Below is a list of potential use cases.

- Mobility enhancements
- Beam management
- Power saving (e.g. choosing DRX parameters)
- Scheduling enhancements
- Autoencoders
- Positioning
- (etc)

30 – Fraunhofer IIS

We think that the SI should focus on a small number of use cases to study the feasibility of AI/ML approaches for the air interface where clear performance gain is observed. We see the following relevant use cases for AI/ML investigations:

- Positioning focusing on LOS/OLOS/NLOS classification and multipath mitigation.
- Beam management focusing on beam prediction

31 – MediaTek Inc.

We would like to start with following use cases in WGs

- AI assisted CSI feedback (R1)
- AI assisted mobility (R2)

32 – Ericsson LM

We see that "AI-enhanced CSI reporting" is a promising area for AI/ML based PHY that involves both gNB and UE and thus excites many different and general issues that needs to be clarified. Given limited time unit allocations and the uncertainties of this topic, we propose to focus on this one example.

If RAN1 discuss this example use case, it can provide learnings for how "AI/ML on PHY" framework might be designed in able to extend this to other use cases.

33 – AT&T

Key use cases

- Mobility
- MU-MIMO
- Beam-management
- Positioning / localization
- App-specific PHY optimization

34 – Orange

We are supportive of a Rel-18 SI on AI/ML focused on channel equalisation, beam selection and link adaptation

35 – Verizon UK Ltd

Priority use cases need to be agreed in the SI based on the expected gains and specification impact. Some important use cases include:

1) MU-MIMO

2) Beam management

3) Positioning / localization

1.3 AI/ML for Air Interface: KPIs and Evaluation Methodology

Companies views:

Feedback Form 3: AI/ML for Air Interface: KPIs and Evaluation Methodology

1 – vivo Mobile Communication Co.

We believe KPI and evaluation methodology should be defined per each area or use case.

For each of the area/use case, we believe the following issue should be discussed:

- Data set for training and test construction: from our initial observation, it can be constructed based on statistical model in 38.901. Data set constructed with 38.901 could effectively emulate local radio environment and statistical model in 38.901 is representative for various practical radio environment; Large number of samples can be generated for training, testing and verification with 38.901;
- Generalization performance evaluation: verification data set construction should consider this aspect. , e.g.,
  large scale parameter perturbation when generating verification data set; system level model in 38.901 is more random and thus should mainly be considered for verification;
- Model alignment between companies: fixed model can be considered during initial calibration between companies; For evaluation purpose, models are selected for applied areas based on company input; more models can also be evaluated by companies and reported further (including model structure and parameters);

2 – Futurewei

· Futurewei believes common evaluation methodologies are very important for comparable comparison and performance analyses.

· We suggest the following aspects to be considered:

a) Dataset construction: This can be from simulation based on common set of configurations, or from actual field data. We suggest using multiple datasets for each identified use case. For comparable comparison across approaches and AI/ML models, we suggest establishing common datasets to facilitate the evaluation effort.

b) Evaluation methodology and KPIs: The evaluation of AI/ML-based approach and KPIs depends on the type of task the AI/ML function is to perform, e.g., regression, classification, or iterative optimization. Futurewei suggests study common methodology and associated KPIs for different AI/ML task type separately.

3 – BT plc

We believe that field data should complement simulated datasets.

4 – Rakuten Mobile

We agree with Vivo that evaluation KPI's and Data set should be defined for each use case, or each use case should be mapped to certain data set/model category.

We propose companies to discuss following topics with AI Data Model and Evaluations.

I. Investigate Potential Data sets for different scenarios of each use case.

II. Data generation, usage and collection methods.

III. For each Scenario/User case agree on using Global data sets OR custom data sets.

IV. Data Granularity and size requirements

V. Conformance specs for each Use Case/Scenario (Minimum Performance improvement)

VI. Interoperability requirements (UE AI Capability, gNB Categories based on AI Model Complexity)

5 – Apple Computer Trading Co. Ltd

Evaluation methodology and KPI can be discussed per use case including:

- Data set: simulated data set per use case can be considered. For example, link level channel model can be used for channel estimation use case, while data set simulated by system level simulation is needed for beam management.

- Testing set and training set: whether test set are generated from the same statistical model as training set. This is to validate the robustness of AI model for various wireless environment.

- KPI can also be discussed per use case. For example, for channel estimation, BLER performance can be used for evaluation comparing to traditional channel estimation method. For CSI, NMSE can be used. For beam management, the metric can be the accuracy for the BSS (beam search space), where a correct BSS is defined as a BSS that includes the best beam.

6 – Qualcomm CDMA Technologies

3GPP needs to agree on datasets for AI/ML study. We think that for the aforementioned use cases in section 1.2, 3GPP can leverage the channel models specified in 38.901 augmented by real-world network logged data to construct datasets for the study. The input and output of an AI/ML model should also be discussed and, if needed, aligned.

From the performance point of view, non-trivial non-ML baseline should be discussed to properly judge the merit of AI/ML-based approaches requiring air-interface impact.

For KPIs, we can define use-case specific KPIs more directly related to the ML model inference performance, such as normalized MSE (NMSE) for cross-node CSF, as well as end-to-end system level metrics such as link or system level throughput.

Model complexity is also an important KPI to consider. It is important to note, however, that complexity is highly dependent on implementation. Nonetheless, an effort to quantify complexity for what is thought a plausible implementation of a given algorithm should be carefully examined and compared against the corresponding non-ML baseline.

7 – SHARP Corporation

Our view is that KPIs and evaluation methodology should be performed per use case or per use cases that have the same characteristics from evaluation aspect, since different use cases may have very different KPIs to be evaluated. For fair and understandable comparisons, common KPIs and evaluation methodology among companies should be defined and captured in 3GPP documents. Considering efforts for determining the KPIs and evaluation methodology for each use case, it is important to narrow down the use cases to be studied.

8 – ZTE Corporation

The general principle of evaluation methodology shall follow 3GPP general framework, including system-level or link-level based on the target use case.

- KPIs should be typical metrics we used in 3GPP, and can be different for different use cases, including throughput, channel estimation MSE, positioning accuracy, BLER, etc..
- Data set is generated based on 3GPP channel model in TR38.901, but parameter randomization shall be taken into account in training and inference. System-level channel model is recommended for both SLS and LLS.
- Assumption for training procedure should be discussed or reported by companies. Offline training is more recommended at early state in the SI.
- Used AI model also needs to be reported by companies. One or a small set of reference models can be aligned among companies at least for calibration.

9 – Xiaomi Communications

Xiaomi

- The study of evaluation methodology should include the following aspects
  - Study the Data set construction including the data collection method, the data format and data set division (e.g., training set, test set) principle
  - Study the Characterization of AI model / algorithm/ training
  - Study Baseline for comparison and calibration is needed
- KPI: depending on different use case, the KPI may be different. But these following common KPI should be taken into account
  - Impact on the device complexity
  - Impact on the power consumption

10 – Intel Technology India Pvt Ltd

- RAN1 should define methodology in order to evaluate robustness of ML solutions (e.g., minimum performance guarantee) for the use-cases indicated above
- Some positioning problems for coordinate inference may require channel model enhancements for spatial consistency (statistical models are not suitable), we think this may not be feasible within the work scope and this issue should be considered when defining use-cases

11 – CATT

KPIs and evaluation methodology shall be studied. Especially, the method of constructing data set for training and testing shall be studied. This is very different from the study we have conducted in the past. Data set plays a key role in the study.

12 – Nokia Corporation

We see the following aspects important when defining evaluation methodology and KPIs.

- Definition of applicable AI/ML framework, training requirements, overhead of training (e.g. online v/s offline), complexity of training and deployment
- Comparability of results (incl. stability and reliability and simulator calibration/repeatability),
- In our view the current 3GPP NR simulation methodology should be extended to enable benchmarking of AI/ML based RAN enhancements, including assumptions for training/exploration and introduction of potential AI/ML-specific KPIs that makes it possible to monitor convergence and stability of such schemes
- Evaluation and KPS to ensure and verify through UE requirements and test cases predictable UE behaviour with AI/ML, and that UEs with AI/ML meet all the existing UE minimum requirements while also obtaining improvements in performance and minimum requirements with AI/ML.
- Extensibility in a multi-vendor environment, i.e., study means to enable feature-independent AI/ML signalling and workflow used by proprietary ML implementations

13 – NTT DOCOMO INC.

At the moment, we have following views

□Data set

A well-constructed dataset should be collected from target environments with good quantity and enough randomness of the parameters. We think diversified raw data can be collected from TS38.901 channel models where a plenty of environments are covered by abundant well-defined wireless channel models.

□KPIs and evaluation methodology

KPIs and evaluation methodology should be defined per use case. The detail of these aspects should be discussed after selecting use case in Rel-18.

14 – Deutsche Telekom AG

The selection of suitable KPIs depends on finally selected use cases. Therefore, this can be done in a second step after use case definition.

The evaluation can be performed via link and system level simulations based on use cases' needs. As the AI/ML models to be used are not intended to be standardized, just the input/output data required, suitable reference scenarios with related KPIs have to be defined that allow comparing the benefits of implemented AI/ML approaches. In addition to performance KPIs also implementation-specific KPIs have to be defined that may take into account the complexity and efforts to train and run AI/ML models (amount of data collection, model transfer/update, etc.).

15 – CAICT

In general, the evaluation methodology and KPIs should be defined case by case. The existing link level and system level simulation could be used for performance verification. It is a basic requirement for further specification works that AI/ML based algorithm can obtain better performance than deterministic algorithm in typical scenario.

Some key issues related to evaluation methodology should be considered:

1 data set for training and test

The first step of performance evaluation for AI/ML algorithm is to establish data set for training and test. Some verification and model exploration works could be based on the data set. Existing 3GPP channel model could be used as baseline.

2 AI/ML model

Although detailed AI/ML model should be left for implementation, some reference AI/ML models could be provided from different companies for performance calibration and further study.

16 – China Mobile Com. Corporation

The study of KPIs and evaluation methodology should include the following issues:

a) Dataset: including the training data and test data. For each use case, all the companies would use the same training data and test data.

b) AI model: For a certain use case, a typical model should be identified, at least for calibration. Companies could use different AI models, but some key common parts of these AI models may be identified, and the typical model can be extracted from the common parts. It is expected that the typical AI model can provided the basic performance.

c) KPI: evaluation metrics could include throughput increase, overhead reduction, complexity reduction, etc.

17 – Spreadtrum Communications

Regarding KPIs, in our mind, other than the conventional performance indicator, the complexity of AI/ML model/algorithm should also be considered.

Regarding methodology, we could take the existed 3GPP methodology as baseline, to trigger the related discussion on methodology.

18 – Guangdong OPPO Mobile Telecom.

From our perspective, the KPIs used for evaluating traditional solutions can be reused as starting point for the study item. For example, CSI feedback accuracy (e.g. MSE or cosine similarity), throughput and overhead are still the suitable KPIs for evaluating AI/ML-based CSI feedback enhancement solutions. Positioning accuracy is used to evaluate AI/ML positioning approaches. One additional aspect is the evaluation of computation complexity of AI/ML inference, we provide our views on this aspect below.

Following study areas can be addressed:

(1) Study and identify the training data set for evaluation (for training, validation and testing)

- Identify adequate amount of training data public to all companies. Or, at least the method to generate the data set needs to be clearly defined and calibrated.
- Examine the quality of the data set, e.g. generalization characteristics. Both single-scenario and multi-scenario data sets are needed.

- Use the computer-generated data set based on TR 38.901 and TR 38.857 (for positioning) as baseline.
- Data set needs to be separately identified for different use cases, e.g. for CSI and positioning respectively.

(2) Study and identify the evaluation process and metrics for identifying performance of AI/ML approaches

- First identify a reference model for each use case.
- Companies can train and optimize their AI/ML models starting from the reference model and provide inference results after training.Necessary information about the model needs to be submitted together with the results, e.g. type of neural network, layer structure, training method, hyper-parameter and loss function.

(3)Complexity analysis for the AI/ML models including investigating the typical computation power on UE and gNB side

(4)Study and evaluate the performance of scenario-fitting models and multi-model switching approach

(5)Study and evaluate the performance of online training of AI/ML models

19 – NEC Corporation

We echo BT's comments that results obtained by using purely simulated dataset may not be convincing. In addition, the cost of the computation resource and power consumption should be taken into account when considering the benefits that the AI/ML will bring.

20 – Samsung R&D Institute UK

To study at least the following:

l How to generate data-set

l Channel model for evaluation

l Evaluation KPI.

21 – Fujitsu Limited

KPIs and methodology can be discussed per use case.

Besides, for a data driven technique, data set construction is the critical issue for AI study. We suggests to have common data sets for train, validation, and test of AI models. Furthermore, AI generalization capability and robustness should be studied.

22 – Lenovo (Beijing) Ltd

The KPIs in this study are use-case specific.

1). For the use case of AI/ML-based CSI compression, the feedback overhead, complexity and throughput with recovered CSI should be evaluated in comparison to the specified techniques in NR Rel-15/16/17.

2). For beam-management, at least overhead, latency, beam tracking/refinement accuracy should be evaluated.

Furthermore, for any AI/ML-assisted approach, the KPIs should also include: a). complexity of the model (e.g., MACs and parameters); b). training overhead (e.g., offline or online); c). generalization issues.

On the evaluation methodology, we suggest at least from calibration perspective:

1). Use a common AI model for the selected use cases without detailed optimization, and the additional complexity from the model can be independently evaluated;

2). Construct the data set for training based on the defined 3GPP channel models with various typical scenarios;

3). Define and evaluate the baseline performance with the common understanding on the system parameter assumptions and related algorithms.

23 – HUAWEI TECHNOLOGIES Co. Ltd.

In general, KPIs should depend on use cases. Some common KPIs for AI/ML schemes can be considered, such as complexity of model training and model inference, model size, signaling overhead, generalization performance, and so on.

For evaluation methodology, the inputs for model training and model inference should be aligned for evaluation and performance comparison. This needs to be discussed and decided at the beginning of the Rel-18 PHY AI/ML study item.

24 – VODAFONE Group Plc

As suggested by BT and NEC, the gathering of real life data sets may well be important.

25 – Ericsson LM

KPIs and evaluation methodologies will need to be defined per any new AI/ML use case.

For AI-enhanced CSI reporting which is the use case Ericsson suggest being studied in Rel.17, the KPIs and methodologies are:

- Use system-level evaluations using NR channel models and scenarios with common understanding of training, validation, and test sets.
  - *Datasets:* We can use 38.901 to generate appropriate synthetic datasets for training, validation, and testing of AI/ML solutions. These datasets can be generated by making appropriate selections/perturbations of parameters. Proprietary datasets can additionally be used within each company, to verify real world performances (like is also done for current CSI reporting features standardized by 3GPP).
- Study achievable DL spectral efficiency of AI-based CSI reporting solutions using Rel 17 features (with a state of the art UE with respect to supported UE capabilities) as the baseline (incl. DL/UL overhead and realistic modeling).
- Study methods, interfaces, and signaling to enable the implementation of proprietary AI-based solutions in the UE and in the network.
  - This requires a common understanding of basic UE and network capabilities for ML model provisioning, training, inference, and updates.
- Study scalability of AI-based solutions w.r.t. input feature sizes/dims that change with system configurable parameters.
- Study feasibility of AI-based solutions in terms of algorithm complexity and energy consumption
- Align with RAN3 higher-level principles (e.g., proprietary ML models are deployed in UE and Network).

26 – Sony Corporation

This is related to the use cases, so we should discuss the details of KPIs and evaluation methodology for each use case in a new study item of AI/ML for air interface.

The KPIs should include the associated signalling and procedures for supporting the AI/ML algorithms.

The KPIs can also consider the reduction in signalling that is possible when an AI/ML approach is taken (e.g. the study can consider whether / how much mobility signalling is reduced for the mobility enhancements use case).

27 – Ericsson LM

KPIs and evaluation methodologies will need to be defined per any new AI/ML use case.

For AI-enhanced CSI reporting which is the use case Ericsson suggest being studied in Rel.17, the KPIs and methodologies are:

- Use system-level evaluations using NR channel models and scenarios with common understanding of training, validation, and test sets.
- *Datasets:* We can use 38.901 to generate appropriate synthetic datasets for training, validation, and testing of AI/ML solutions. These datasets can be generated by making appropriate selections/perturbations of parameters. Proprietary datasets can additionally be used within each company, to verify real world performances (like is also done for current CSI reporting features standardized by 3GPP).
- Study achievable DL spectral efficiency of AI-based CSI reporting solutions using Rel 17 features (with a state of the art UE with respect to supported UE capabilities) as the baseline (incl. DL/UL overhead and realistic modeling).

28 – AT&T

Emphasis should initially be on ML/AI algorithms embedded in functionality of the RAN to enable **real-time**decisions with distributed data. How the UE may be able to assist and or actively participate in the optimization/decision making framework can also be studied.

29 – Ericsson LM

KPIs and evaluation methodologies will need to be defined per any new AI/ML use case.

For AI-enhanced CSI reporting which is the use case Ericsson suggest being studied in Rel.17, the KPIs and methodologies are:

- Use system-level evaluations using NR channel models and scenarios with common understanding of training, validation, and test sets.
- *Datasets:* We can use 38.901 to generate appropriate synthetic datasets for training, validation, and testing of AI/ML solutions. These datasets can be generated by making appropriate selections/perturbations of parameters. Proprietary datasets can additionally be used within each company, to verify real world performances (like is also done for current CSI reporting features standardized by 3GPP).
- Study achievable DL spectral efficiency of AI-based CSI reporting solutions using Rel 17 features (with a state of the art UE with respect to supported UE capabilities) as the baseline (incl. DL/UL overhead and realistic modeling).

- Study methods, interfaces, and signaling to enable the implementation of proprietary AI-based solutions in the UE and in the network.
- This requires a common understanding of basic UE and network capabilities for ML model provisioning, training, inference, and updates.
- Study scalability of AI-based solutions w.r.t. input feature sizes/dims that change with system configurable parameters.
- Study feasibility of AI-based solutions in terms of algorithm complexity and energy consumption
- Align with RAN3 higher-level principles (e.g., proprietary ML models are deployed in UE and Network).

30 – MediaTek Inc.

On dataset - common framework, but individual dataset for each use case

- Simulated and real channel data can both be used for evaluation
- also need to establish test dataset to verify the performance, such test dataset needs to reflect the real world imperfection

On KPI - individual for each use case

- radio performance related KPI can be used per use case
- other AI/ML algorithm related KPI can be added

31 – T-Mobile USA Inc.

AI/ML needs to be a SI in R18, SA& RAN need to understand the potential use cases, their benefit relative to current techniques and more importantly if AI/ML algorithms will be prescribed in the specification or will 3GPP follow the RF scheduler model were 3GPP defines the API's needed for AI/ML implementations.

Privacy and security are important aspects of AI/ML, these need to a integral part of any AI/ML specification efforts.

1.4 AI/ML for Air Interface: Network and UE involvement

Companies views:

Feedback Form 4: AI/ML for Air Interface: Network and UE involvement

1 – vivo Mobile Communication Co.

Dependent on targeted category of use cases, involvement of network and UE may be different.

For Cat1 (as shown below), network and UE involvement is transparent to each other;

For Cat2 (as shown below), UE can be involved for some use cases; for example, new RS pattern may be designed based on AI based channel estimation; network can be involved for some use cases, for example, channel state prediction based on UE report.

For Cat3 (as shown below), network and UE involvement is necessity. Both sides are aware of the AI operation and provide assistance information to the other side for training or for further inference.

For Cat4 (as shown below), UE involvement is necessity. The expected operation is based on joint inference at network side and UE side.

Cat1 Cat 4 refers to the following:

- Cat1: AI/ML related training and inference are all conducted at one side of network or UE and is transparent to the other side;
- Cat2: AI/ML related training and inference are conducted at one side of network or UE, but requires additional signaling or procedure enhancements between two sides, potentially with existing signaling framework. Additional information is not directly related to training and inference, e.g., capability, new patterns etc.; Training or inference is done by implementation and not explicitly seen in the specification
- Cat3: AI/ML related inference is conducted at one side of network or UE, with assisted training information exchanged between two sides; E.g., neural network models provided as a reference for UE or for aggregation
- Cat4: AI/ML related inference are conducted together at both sides of network and UE training maybe conducted at one side or both. Information related to inference need to be exchanged between both sides;

2 – Futurewei

· Depending on the use case, some use cases may involve collaboration between network and UE in AI/ML operation. The pros and cons including overhead introduced and inter-operability need to be carefully investigated.

· Futurewei suggests establishing high-level principles and framework in handling the network and UE collaboration for AI/ML-based approach at PHY layer via use case study.

3 – InterDigital

Extent of NW and UE involvement is use case specific. We think that one outcome of initial study phase is the assumption on the applicable deployment(s) for the identified use case (i.e. whether training is offline and/or online, whether inference is performed by UE, network or both)

4 – Rakuten Mobile

Agree with Futurewei and InterDigital- This should be based on each Use chase

5 – Apple Computer Trading Co. Ltd

It is important to identify the role of Network and UE in each use case:

1. Whether the training is done at network, or at UE. The jointing training such as federated learning is proposed to be de-prioritized for R18 study.

2. Whether the inferencing is done at network, or at UE. The distributed inferencing is proposed to be de-prioritized R18 study.

3. Once the benefit of AI based use case is identified and agreed, assisted signaling can be further discussed.

6 – Qualcomm CDMA Technologies

As discussed, we foresee different levels of inter-node collaboration for which the exchange of ”information” will be fundamentally different.

For each use case and inter-node collaboration level, network and UE involvement should be discussed. This includes, e.g., signaling and measurement procedures to provide training data for online training, signaling and measurement procedures to provide input to ML model for inference, signaling and/or feedback of output from ML inference, and AI/ML model management such as model provisioning, configuration, and activation/deactivation for inference and training, and model download and upload.

Possibility of testing a given performance metric should be considered and extensions of current RAN4 procedures, if and when needed, should be carefully examined.

7 – SHARP Corporation

Our view is that, as the starting point, it is better to focus on the simplest form of AI/ML applications such that the model training and the inference are performed at the network side. Depending on the use case, enhancements on RS configuration/measurement reporting/signaling enhancement may or may not be required. The next step could be model training at the network side and transferring the inference model to the UE. The UE performs inference using the indicated model. Considering the complexity and specification impact issues, we believe step-by-step methodology could be efficient from the specification work point of view.

8 – ZTE Corporation

The following aspects need to be studied in the SI about NW and UE involvement. Different aspects may apply on different use cases.

- UE reporting information to assist NW's AI operation, e.g., for beam/mobility prediction, positioning, UL RS overhead reduction, etc.
- NW's indication to assist UE's AI operation, e.g., for DL RS overhead reduction, etc.
- NW's indication based on AI operation in NW side, e.g., for UL RS overhead reduction, beam/mobility prediction, etc.
- NW and UE's collaborative AI operation or split inference involving AI model alignment, e.g., for CSI compression feedback

9 – Xiaomi Communications

UE specific information, e.g. UE location, is very beneficial to improve inference accuracy. However, due to privacy requirements, these information may not be available at gNB. UE based AI inference could utilize these UE information without privacy issue. UE collects these information and ran AI inference locally. The inference result could be used locally or reported to NW, which has no privacy issue. On the other hand, NW could also be involved from the functionality control and AI model transmission perspective. Therefore, we believe both NW and UE impact should be studied.

10 – Intel Technology India Pvt Ltd

We think model transfer across the air-interface leads to higher specification impact and should be properly justified. We prefer to limit use cases that may require model transfer across the air-interface (for example CSI use-case).

11 – CATT

Our preference is to avoid splitting the training procedure between network and UE. If the training of the AI model is split between network and UE, then network and UE need to synchronize the parameters/intermediate results of AI module leading to complex signaling exchange between network and UE. It would be very challenge to finish this kind of study in Rel-18.

12 – Nokia Corporation

In order to make the introduction of AI/ML successful for radio air interface we see that it is important in the early phase to study how to ensure and verify predictable UE behaviour with AI/ML and that the UE utilizing AI/ML for air interface meets all the existing UE minimum requirements while also obtains improvements in performance and minimum requirements with help of AI/ML. Therefore, early studies of UE requirements and test cases for the AI/ML use cases where UE utilizes AI/ML.

13 – NTT DOCOMO INC.

Collaborated ML models on NW and UE require coordinating signaling between UE and NW. We prefer to avoid use cases involved in these signaling due to the large specification impacts. Use cases with AI/ML deployment on NW is of our high interest as a first step.

14 – Deutsche Telekom AG

Different learning methods may require different approaches w.r.t. AI/ML implementation. This has to be considered for the use cases under discussion as this may impact how network and UE are working together.

In addition, also the demand for offline and online training has to be clarified per use case. Our preference would be to have initial offline AI/ML model training (incl. validation and testing) outside of the RAN, e.g. in the OAM domain, and only fully validated/tested models are then transferred to RAN and/or UE. The case that online model training inside RAN and/or UE during operation has to be considered is again dependent on the use cases finally selected.

We would prefer to initially focus on simple use cases to avoid complex model management discussions.

15 – CAICT

For different use cases, AI/ML model deployment will not be the same. Network and UE involvement for AI/ML enhancements includes AI/ML model training, inference, update and transferring, data collection from UE and gNB side and assistant control information delivering. Some guidelines for network and UE involvement could be provided in use case identification phase and detail discussions could be given based on performance evaluation.

16 – China Mobile Com. Corporation

We should identify whether online training or AI mode update is needed, if yes, how to train/update the AI model should be studied. It may include model offloading or collaboration between gNB and UE.

17 – Spreadtrum Communications

It depends on specific use cases. Some use cases only need network involvement, e.g., UL positioning; some use cases only need UE involvement, e.g., RS overhead reduction; some use cases need both NW and UE involvement, e.g., CSI feedback with low overhead. Given that network and UE involvement at least could reflect the workload, it can be one criteria for down-selection among use cases. With what we have said, we are open for the use cases.

18 – Guangdong OPPO Mobile Telecom.

The following network and UE involvements can be studied by investigating the procedures:

- The physical procedure on 5G network and UE side facilitating AI/ML model inference, e.g. input/output format of inference data, inference parameter configurations
    - Different procedures can be studied for different solutions for a use case (e.g. different positioning solutions).
    - Model switching/downloading is also considered.
- If online training is considered, the physical procedure on 5G network and UE side facilitating AI/ML model training, e.g. input/output format of training data, training parameter configurations.
- Measurement/reporting enhancement for AI/ML-based approaches.
- Timeline enhancement for AI/ML-based approaches.

19 – NEC Corporation

We share similar view that it relates to the use cases. Network, UE, or both could be involved. So in the SI phase, we don't want to preclude anything.

20 – Samsung R&D Institute UK

To study could include AI framework, including:

- Necessity of new signaling and measurement procedures
- Appropriate AI framework (inc. separate or Joint ML operation between UEs and NW)
    - Can improve performance in terms of latency and energy efficiency of training/inference
    - Can offload computation-intensive and energy-intensive parts to network endpoints

21 – Fujitsu Limited

It is a use case specific issue. We support to clarify the assumptions of network/UE involvement in training/inference when concluding use case.

22 – Lenovo (Beijing) Ltd

To have a general and extendable study, it is better to have a uniform framework to support flexible involvement schemes, including UE-only, network-only and joint network and UE. When considering joint AI/ML application at both sides, the NG-RAN for AI/ML could be jointly studied in terms of complexity/overhead related to the model transferring over the air interface.

23 – HUAWEI TECHNOLOGIES Co. Ltd.

Network and UE involvement should be discussed for each specific use case, by considering the procedures of model training and model inference. There can be different ways of network and UE involvement. The model training and model inference can be applied at network side, UE side, or both, depending on the use cases. The signaling between network and UE can also be different for each use case. The corresponding details need to be ironed out during the study item.

24 – Ericsson LM

An important part of the SI is to determine the level of network and UE involvement for different use cases.

25 – Sony Corporation

The study should consider how the AI/ML is trained and which entity trains the AI/ML model. Training of the AI/ML model can be executed at both/either network and/or UE. This also should be studied in new study item of AI/ML for air interface.

The study should also consider where inference of the AI/ML model is performed.

26 – Ericsson LM

An important part of the SI is to determine the level of network and UE involvement for different use cases.

27 – AT&T

Emphasis should initially be on ML/AI algorithms embedded in functionality of the RAN to enable **real-time**decisions with distributed data. How the UE may be able to assist and or actively participate in the optimization/decision making framework can also be studied.

28 – MediaTek Inc.

For each use case

- Study UE and network AI/ML function split
- Study required data/signaling exchange on physical/logical channel for data collection and model delivery
- Study AI/ML capability awareness among UE and NW nodes

1.5 AI/ML for NG-RAN

Companies views:

Feedback Form 5: AI/ML for NG-RAN

1 – vivo Mobile Communication Co.

AI/ML for NG-RAN should avoid impacting UE.

2 – Futurewei

· For NG-RAN use cases, Futurewei suggests leveraging the framework and study that have already been established under Rel-17 RAN3 SI: FS_NR_ENDC_*data*_collect.

· Any changes or evolution required should be identified and coordinated.

3 – InterDigital

NG-RAN work should be done in two parts. The first part is a work item in support of the current study item use cases. This work item should not have AI functionality in the UE but potentially support new UE functionality like new measurement support. The second part should be a study for new NG-RAN use cases beyond those in the current study and could include some with AI functionality in the UE.

4 – Apple Computer Trading Co. Ltd

Out of the 3 prioritized use cases in RAN3 led SI: energy saving, load balancing and mobility optimization, we support the use case "Mobility performance enhancement, including power consumption reduction and latency reduction" from RAN2 point. Further clarify is needed to have a clear definition of "performance" and "power". Reducing handover failures and optimizing CHO, e.g. in terms of network resources reserved in candidate target cells, can be part of the enhancement for this use case.

5 – Qualcomm CDMA Technologies

The conclusion of R17 AI/ML study by RAN3 should be used as base for R18 AI/ML work item. It is expected that the R17 conclusion would include at least:

- Data collection enhancements for the three use cases: network energy saving, load balancing and mobility optimization
- Inter-node information enhancements for AI/ML model output of the three use cases
- Performance monitoring of AI/ML based solutions

Depending on the study phase conclusion, the R18 work item may also include:

- Model management procedures, e.g. model configuration, model update, model activation/deactivation, model training, model/analytics querying
- Data management procedures, e.g. data request/subscription, data/event reporting, data querying.

If not concluded in R17 study, these should be included in R18 study item.

In addition, R18 should study architecture and procedure enhancements for AI/ML, including

- Network entities and interfaces for model training/inference, data management, model management
- Procedures for online training and federated learning
- Procedures for inference configuration in UE and network
- Procedures for data management and model management
- Mobility support for AI/ML based solutions
- UE assistance for AI/ML configuration.

R18 should continue the use case study. If possible, some standardized neural network functions like Y = F(X) can be defined based on the use case study. The AI/ML model supporting the neural network function should be up to implementation.

6 – SHARP Corporation

AI/ML for NG-RAN use cases in Rel-18 could be extended based the existing SON/MDT specifications and the Rel-17 RAN3 study item.

7 – Intel Technology India Pvt Ltd

For Rel-18, we think AI/ML for NG-RAN can be separated into two working streams:

- one WI: normative work of "FS *NR*ENDC*data*collection", including signalling exchange via XnAP/F1AP/NGAP/etc of load balancing, network energy saving and mobility optimization.

- one SI: a continuous study item on 1) deployment scenario and functional framework evolution; 2) supporting new use cases which are deprioritized in Rel-17, such as Urllc, network slicing, QoE optimization, IAB, etc; 3) study standard impact of supporting distributed learning across NG-RAN/federated learning between NG-RAN node and other network nodes

For any impact of UE (e.g. additional assistance information), RAN2 should be involved in the study or WI.

8 – ZTE Corporation

We see the necessity of normative AI/ML for NG-RAN based on output from study stage. Three AI based use cases have been highly prioritized and AI functional framework is going to be stable. Therefore, we support the normative work of SI/ML for NG-RAN in Rel-18 with following objectives as start:

- Specify solutions, measurements(input/output), signaling, and procedures for Network Energy saving;
- Specify solutions, measurements(input/output), signaling, and procedures for Load Balancing;
- Specify solutions, measurements(input/output), signaling, and procedures for Mobility Optimization;
- Interface function enhancements to support AI/ML functionality in both non-split architecture and split architecture, the common interface functions for AI should be studied;
- Signaling impacts to support transferring large size of AI/ML data in a security way.

In addition, we see the beneficial to further study other user cases and other aspects of AI/ML for NG-RAN based on the output of study stage. We support 1 separate Study item in RAN3 or merged the following objectives into AI/ML for Air Interface study led by RAN1.

- Study and identify other high-layer AI/ML based use cases (e.g., QoE optimization, Network Slicing, etc);
- Support AI/ML related information or AI/ML related model exchanged between 5GC and RAN when needed.

9 – CATT

We agree with others that Rel-18 work on high layer AI should include 2 parts.One is WI which aims to do the nomative work for the use cases that had already been fully discussed in Rel-17,e.g. Energy saving,Load balancing,Mobilty management.Depending on the outcome of Rel-17 SI,the corresponding WI may involve both RAN2 and RAN3 if the necessity on the information transfer on network interface and UU interface is justified.The other part is Rel-18 SI in which further study on other high layer realetd use cases could be pursued.

10 – CAICT

AI/ML for NG-RAN specification works should be based on the output of R17 study item.

11 – Deutsche Telekom AG

The work on AI/ML for NG-RAN can be split into 2 parts.

The first one is a WI based on the outcome of the Rel-17 SI performed in RAN3. This WI should consider the normative specification of the relevant parts of the functional framework for RAN intelligence within NG-RAN as described in TR 38.817 and the interface impacts of the use cases covered in that TR.

The impact of further use cases (e.g. mobility optimization on DAPS/CHO level, coverage/capacity optimization especially for Massive MIMO antenna usage) can be studied in a new SI. In such SI, we propose to additionally study and identify the possibilities on introduction of a AI/ML framework on Service Based Architecture (SBA) principles in the NG-RAN, which are already used in 5GC and OAM, and to enable a soft migration into a fully cloudified environment for the entire Data Analytics and ML/AI incorporating NWDAF and MDAS approaches of SA2 and SA5. This should especially cover a lifecycle management process for AI/ML modes aligned across different 5GS domains.

12 – China Mobile Com. Corporation

R18 AI/ML work item should follow the conclusion of R17 AI/ML study lead by RAN3. The detailed scope should not be discussed at RAN93e-meeting, since the study is still ongoing.

Further study on other use case is also expected in Rel-18.

13 – TELECOM ITALIA S.p.A.

as stated in the general part, we share the view expressed by Deutsche Telekom.

An initial Work Item to ensure that data can be collected followed by a long term study for a specification phase in Rel 19 is our preferred way forward.

To ensure data collection (Work Item phase), features like MDT should be made mandatory in the UE and proper interfaces / data formats should be specified from the network side.

14 – Spreadtrum Communications

We prefer to directly kick off nominal work on NG-RAN, based on Rel-17 study work. But if new use cases can be identified, we are also fine to have related study.

15 – Guangdong OPPO Mobile Telecom.

AI/ML has been studied by SA1 in SP-191040, the output is captured in 22.874. From RAN perspective, SA1 AMMT requirements can't be met by current 5G system, so we propose to study the following:

- Study the potential gaps between 5G NR performance and the requirements defined by SA1 AMMT requirements:
  - High UL experienced data rate e.g. few hundred Mbit/s;
  - Low latency high reliability in combination with high UL experienced data rate;
  - support 'aggregated performance' for a group of UEs where the worst performing member defines the performance of the entire group;

In addition a follow-up WI can be led by RAN3 to complete the normative work in TR37.817. Three use cases can be considered, i.e. Power saving, Load balance and Mobility optimization.

16 – NEC Corporation

We propose to consider two directions in Release 18:

- Normative work based on the results of Release 17 RAN3 study
- Study on enhancements of NG-RAN AI/ML.

Normative work based on the results of Release 17 RAN3 study could include the following:

- Focus on the current NG-RAN architecture and interfaces
- Focus on the use cases prioritized in Release-17 RAN3 study
- AI/ML Model Training and Model Inference are black box
- For the prioritized use cases, several deployment options identified in Release-17 RAN3 study should be considered
- Specify inputs to Model Training and Model Inference from Data Collection
- Specify outputs from Model Inference to Actor and Data Collection
- Existing SON procedures and may serve as a basis and may be further enhanced
- New RAN AI/ML specific procedures may be defined
- Existing measurements may be reused
- New specific measurements may be defined

Study on enhancements of NG-RAN AI/ML could include the following:

- Non-prioritized use cases from Release-17 RAN3 study and new use cases
- New interfaces and/or functional entities may be considered if needed
- Multi-vendor interoperability between different AI/ML model parts (Data Preparation, Model Training, Model Inference)
- Possibility to pre-select and pre-train some ML models for given tasks could be studied
- Integration and collaboration of OAM AI/ML, 5GC AI/ML, NG-RAN AI/ML, and Air Interface AI/ML could be considered, as well as, OAM management aspects of NG-RAN AI/ML

17 – Samsung R&D Institute UK

Based on the outcome of the RAN3 study in Rel-17 RAN3 SI: FS*NR*ENDC*data*collect, the AI-based load balancing, energy saving and mobility optimization need to be supported by normative specifications in Rel-18.

Any new use cases should be studied firstly e.g. in a separate SI.

18 – Nokia

The R18 WI scope should be based on the conclusions of the ongoing R17 SI in RAN3. Based on the current status in RAN3, we expect the following can be specified:

- Architecture options for the prioritized R17 use cases.
- Data collection for Training and Inference phase. This includes requesting data for training/inference of an ML Model from other network entities on a use-case specific basis, and defining associated new measurements (as needed) in the NG-RAN.

In addition, a concurrent R18 SI can continue to study further enhancements to the prioritized R17 use cases and additional new use cases, considering also potential UE involvement e.g. ML algorithm located in the UE.

19 – Fujitsu Limited

AI/ML for NG-RAN can start its normative work based on the results of Release 17 RAN3 study.

20 – KDDI Corporation

See our comments on 1.2 AI/ML for Air Interface: Use cases to focus.

21 – Lenovo (Beijing) Ltd

We suggest the evolution of AI/ML for NG-RAN should include two parts in NR Rel.18:

- A normative work for the study outcomes including the prioritized use cases (e.g., energy saving, load balance, and mobility optimization) following the conclusions of 3GPP RAN3 R17 study item.
- A further study on potential RAN enhancements to support distributed intelligence to fully utilize the AI/ML capability among dense deployed RAN nodes. For example, RAN can study mechanisms to facilitate AI management procedures, AI/ML model distributing/update, computing offloading, data sharing with security and privacy etc. among RAN nodes.

22 – HUAWEI TECHNOLOGIES Co. Ltd.

In general, the scope of Rel-18 AI/ML for NG-RAN should be based on the agreement in R17 SI,

- Existing network architecture should be reused, and new interfaces shall not be introduced.
- Detailed AI/ML algorithms and models shall be left for implementation.
- The study focuses on the analysis of data needed at the model training/inference function from Data Collection, while the aspects of how the model training/inference function uses inputs to derive outputs are out of the scope.

Out of the use cases studied in Rel-17 SI, our current view is that energy saving, load balancing and mobility enhancement may be considered for normative work in Rel-18.

23 – Ericsson LM

For Rel-18, we should take the selective use cases from Rel-17 technical report into normative phase:

- AI/ML mobility optimization
- AI/ML for load balancing
- AI/ML for network energy saving

According to the agreed high-level principles, we should maintain the current RAN architecture and Interfaces.

Model deployment Update/Model Training is implementation dependent and should not be specified.

24 – VODAFONE Group Plc

We are generally supportive of the views from DT and Telecom Italia.

In addition, there may be some RAN work needed to make information gathered in RRC-Inactive available to the 5G Core's NWDAF.

25 – Ericsson LM

For Rel 18, we should take the selective use cases from Rel-17 technical report into normative phase:

- AI/ML mobility optimisation
- AI/ML for load balancing
- AI/ML for network energy saving

According to the agreed high-level principles, we should maintain the current RAN architecture and Interfaces.

Model deployment Update/Model Training is implementation dependent and should not be specified.

26 – Sony Corporation

Work on AI/ML for NG-RAN should be based on the outcome of the RAN3 study.

27 – Orange

We are supportive of specifying the required metrics and data collection to support AI/ML algorithms in a number of use case for RRM optimisation: traffic steering, load balancing, mobility and interference management, RAN slicing and energy efficiency.

28 – AT&T

Taking the Rel-17 study item conclusions as a starting point, the work should focus on RAN interface enhancements to support training and model updates for **non-real time** and **real-time** use cases.

29 – MediaTek Inc.

If R3 Rel-17 SI reaches conclusion, it shall go into normative phase in Rel-18. In general, further study can continue to identify new data collection for performance optimization with strong justification and architecture enhancements for training data aggregation among network entities, distributing computing, and efficient model delivery.

30 – Verizon UK Ltd

We see two activities for AI/ML for NG-RAN in Rel-18

1) A WI based on RAN3 Rel-17 SI that includes basic AI/ML framework and initial use cases

2) A follow-up SI that includes further enhancements/refinements to AI/ML framework and additional phase 2 use cases. This should also include RAN interface enhancements to support model training/updates, and potential alignment with O-RAN.

1.6 Other proposals

Companies views:

Feedback Form 6: Other proposals

1 – vivo Mobile Communication Co.

A possible arrangement of the study can be shown as following:

- Initial phase (6 9 months) for use case study includes collection of AI/ML background information and terminology alignment, categorization of use cases that may benefit from AI/ML application on air interface and recommendation of focused categories and use cases for the follow-up WG level study in Rel-18;
- Focused study phase (9 12 months) includes evaluation of performance of applying AI/ML in the identified areas, framework and procedures of applying AI/ML in the identified areas and potential specification impact of applying AI/ML in above areas, including related procedures;
- Work item phase to be started based on the output from the WG study item (Maybe started in Rel-19);

2 – Futurewei

· Futurewei suggests the Rel-18 SI to include the following:

a) Identifying AI/ML-based PHY layer use case(s)

b) Conducting use case study, identifying common evaluation methodology for performance analysis, studying standards impact for each identified use case.

· AI/ML for PHY SI will be studied for the full Rel-18 term.

3 – BT plc

Duration of the study should be discussed following the identification of objectives for AI/ML-enabled use cases. Duration of such use case identification phase could be set for 9 months (assuming a 18-months release).

4 – Qualcomm CDMA Technologies

We believe that is important to consider ML management procedures (RAN2) as part of the scope of the study.

5 – SHARP Corporation

We believe it is important to define a clear and constrained scope in the study item, e.g., determine the most essential and typical use cases, define the corresponding KPIs and evaluation methodology, and identify the potential specification change requirements.

6 – ZTE Corporation

It should be good to align some principles for WG-level study in RAN before Rel-18 starts, for example, prioritizing selected use cases, prioritizing offline training, striving to make AI model as implementation, generating data set based on 3GPP channel model, etc.

7 – CAICT

We support vivo's proposal and two phase study is beneficial for the study of AI/ML for air interface.

8 – TELECOM ITALIA S.p.A.

Approve

- a Work Item based on the outcome of the RAN3 SI and focusing on enablers for data collection
- approve a long term SI on physical layer and NG-RAN aspects

9 – Guangdong OPPO Mobile Telecom.

We prefer to organize the study item in two phases: The first phase is led by RAN, to identify the high-level framework of the study, including the principle of evaluation methodology and studied use cases. Then the detailed study and evaluation are carried out in RAN1 and RAN2

2 Intermediate Round

First off, the moderator would like to thank all the companies for their inputs. In this intermediate round of discussions, the moderator would like to identify a few areas aiming at further understanding the underlying implications of the related Rel-18 candidate projects.

2.1 AI/ML Projects

Companies' inputs have broadly made it clear that we are discussing two distinct but related projects in this Thread. As a result, the moderator suggests capturing the following two candidate projects for Rel-18 with some high level notes for further discussion:

Table 1: AI/ML related candidate projects for Rel-18 [*] No need to further discuss scope in this phase of R18 discussion.

	Temporary title	**SI/WI**	**Primary WG**	**Secondary WG**	**Notes**
1	AI for NG-RAN	WI	RAN3	RAN2	Based on the outcome of RAN3-led Rel-17 SI *FS_NR_ENDC_data_collect*. [*]
2	AI for Air-Interface	SI	RAN1	RAN2, RAN4	SI for entire Rel-18 timeframe focusing on a few identified use cases.

Companies are welcome to share their views about the Table above identifying two candidate projects for Rel-18 related to AI/ML:

Feedback Form 7: AI/ML candidate projects structure

1 – Futurewei

Futurewei supports the identified two candidate AI/ML projects in Rel-18.

2 – vivo Mobile Communication Co.

As commented by several companies, AI for air-interface may need to be split into several phases.

Even if we agree before December the high level areas that need to be focused in Rel-18, each of the high level area may still contain drastically different use cases, for example, as shown in our comments above, there are four categories of use cases in each of the high level areas.

An initial phase for categorization and clarification of these hidden use cases would be helpful for better planning of the study item and also beneficial for long term evolution of this feature.

3 – Convida Wireless

We are supportive of two candidate AI/ML projects for Rel-18 suggested by moderator.

4 – Qualcomm Incorporated

Qualcomm agrees with the two candidate projects for Rel-18.

5 – Apple Hungary Kft.

Agree with the project structure

6 – CATT

Agree with vivo that an initial phase of study is helpful to gain better understanding of AI/ML for air interface and narrow down use cases.

We think AI for NG-RAN may need to have a WI and SI separately, depending on the outcome of Rel-17 SI. The work on WI is to do normative work for use cases which had conclusion in Rel-17. The SI would focus on the study of use cases which could not be covered by the Rel-17 SI.

7 – Beijing Xiaomi Mobile Software

We agree with the proposal in general.

8 – Intel Technology India Pvt Ltd

Support the proposed project structure

9 – Deutsche Telekom AG

We agree with the proposal to set up a SI for AI/ML for Air Interface and to start a WI on AI/ML for NG-RAN based on RAN3's Rel-17 SI output.

In addition to the normative work for NG-RAN we see also the need to set up a new SI to discuss further use cases and possible architectural updates (alignment of AI/ML framework across 5GS domains, use of SBA approaches, ...).

10 – Spreadtrum Communications

We support moderator's proposal to set up one WID for NG-RAN, and one SID for air interface led by RAN1.

11 – Verizon UK Ltd

Agree with DT view. In addition to a AI/ML WI based on Rel-17 RAN3 SI, we also see a need for an AI/ML SI for NG-RAN that studies further architecture improvements and new use cases.

12 – NTT DOCOMO INC.

We support identifying two projects for Rel-18 related to AI/ML.

13 – LG Electronics Inc.

We agree that these two items need to be split.

14 – SHARP Corporation

We support moderator's summary.

15 – CAICT

we also propose to have an initial phase for use case study for AI/ML for air interface. It will be helpful for having a better understanding of candidate use cases and making a long-term evolution for AI/ML based air interface enhancements.

16 – Guangdong OPPO Mobile Telecom.

We are general fine with the project structure. For SI AI for Air-Interface it is not clear why RAN4 should be involved. In addition we wonder whether we should also consider the throughput bottleneck when online delivery of AI/ML model is necessary during the study

17 – Fujitsu Limited

We support this summary.

18 – Fraunhofer IIS

Fraunhofer supports the identified two candidate AI/ML projects for Rel-18.

19 – ZTE Corporation

We are supportive to have these two parallel projects. Three comments on the details

- It seems the remaining sections are all for the RAN1-led SI. For NG-RAN WI, we should also discuss the potential scope. Perhaps the moderator can clarify where to discuss this part in this email thread.
- For the RAN3-led WI, the scope and objectives shall leverage the outcome of the RAN3 SI. In addition, we suggest to discuss other high-layer use cases and other enhancements for AI in NG-RAN in new RAN3-led SI or in the RAN1-led SI, which is not concluded in the current RAN3 SI.
- For the RAN1-led SI, we think it is a bit early to conclude we will use the whole Rel-18 timeframe to conduct the study. If there are some techniques which requires low or minimal specification impact, we can quickly do the normative work in Rel-18 if time allows.

20 – MediaTek Inc.

We support the proposed structure. On R3 WI, we believe the new data collection from UE shall be carefully decided based on well-captured evaluation and the UE impact is similar to SON/MDT with mandatory user consent.

21 – NEC Corporation

We support both projects.

Based on the first round discussion, many companies proposed to have continuation WI and new SI related to AI/ML for NG-RAN.

Would it be possible to clarify whether both WI and SI in RAN3 are included in "1. AI for NG-RAN"?

22 – KDDI Corporation

We support two projects in Rel-18.

23 – Sony Corporation

We agree with the project structure except for the Note of AI for air interface. We think use cases shouldn't be down-selected at this stage. We should have an initial phase of the study to identify use cases.

24 – RadiSys

Agree with the project structure

25 – Nokia Corporation

Moderator's proposal is acceptable for us. However, we believe that RAN3 should continue to study additional AI/ML enhancements and use cases for NG-RAN which do not end up within the scope of the agreed Rel-18 work item. Continuation of the AI/ML for NG-RAN study could be added as a 3rd project (RAN3 primary, RAN2 secondary).

26 – Lenovo (Beijing) Ltd

For the study work, we suggest **one SI for NG-RAN led by RAN3, and one SI for air interface led by RAN1**.

Regarding AI for NG-RAN, these two concerns have been proposed by many companies in the initial round:

- A normative work for the study outcomes including the prioritized use cases (e.g., energy saving, load balance, and mobility optimization) following the conclusions of 3GPP RAN3 R17 study item.

- A further study on potential RAN enhancements to support distributed intelligence to fully utilize the AI/ML capability **among dense deployed RAN nodes**. For example, RAN can study mechanisms to facilitate AI management procedures, AI/ML model distributing/update, computing offloading, data sharing with security and privacy etc. **among RAN nodes**.

Especially, the second part is proposed at least by Lenovo, Intel, NEC, QC, ZTE, DT, Nokia, Vodafone and etc. Because this part is the higher layer issue related with RAN3/2, it is necessary to consider the relationship with the SI for air interface.

There are three options to handle the second study part:

- Option 1: A separate SI for NG-RAN, led by RAN3

- Option 2: Included in the SI for air interface, as a RAN3-led component
- Option 3: Included in the WI for NG-RAN, led by RAN3

In our view, neither Option 2 or 3 seems to be an optimal way, because:

1). For the limited TU budget, the SI for air interface will mainly focus on the use cases selection, KPI and evaluation methodology determination, not much TU for the study on the RAN enhancement with Option 2;

2). It is highly doubted that companies will agree to enhance RAN node (e.g., Xn interface) for the potential AI management issue in the scope of normative work with Option 3.

Therefore, from Lenovo/Motorola Mobility point of view, we prefer Option 1, i.e., a separate SI for NG-RAN is needed for the study on potential RAN enhancements.

27 – HUAWEI TECHNOLOGIES Co. Ltd.

The relationship of the two projects should also be captured in order to ensure consistent framework with compatibility and interoperability between features/projects. At a minimum, the principles from on-going RAN3-led SI should remain applicable for the SI of AI for Air Interface, which are

- Detailed AI/ML algorithms and models are left for implementation;
- User data privacy needs to be preserved;
- Support of AL/ML shall reuse the existing RAN architecture and new interfaces shall not be introduced.

28 – Samsung Electronics Polska

We agree with the project structure in general. We also support to have an independent SI on AI for NG-RAN to study the use cases which could not be covered by the Rel-17 SI.

29 – China Mobile International Ltd

Thanks for moderator to raise the project structure topic, we can have discussion but it seems still to early to make decisions.

In our view, firstly, no matter what projects are planned in Rel-18, it should be consistent with the AI/ML framework studied and defined in Rel-17 data collection enhancement SI. For the project structure, maybe there are 3 components needs to be considered.

1) A WI focusing on AI for NG-RAN, to start the normative work following the conlusions of Rel-17 SI

2) A study focusing on AI for air interface

3) Continuation of the study on AI for NG-RAN, e.g., additional use cases, framework and architecture enhancement, etc. How to handle this study may have several ways, to extend the current one, to have a third project or to combine with the AI for air interface.

However, this can be dicussed later. In the moment, the focus should be convergence of the scope and objectives of AI for air interface. The objectives of the other two can be discussed when the existing SI alomost finishs.

30 – Rakuten Mobile

We agree with proposed project structure

31 – Ericsson LM

Ericsson is supportive of an *AI for Air Interface* SI lead by RAN1, with secondary involvement from RAN2 and RAN4.

Ericsson proposes that the *AI for Air Interface* SI focuses on a single example use case to gather learnings for future ML based work in RAN1. Learnings from this single example use case can be documented in a technical report (TR).

Hence, we suggest to modify the note to "SI for entire Rel-18 timeframe focusing on a single use case for detailed studies, while possibly performing a general study (e.g. without evaluations) on additional identified use cases"

32 – InterDigital

We agree with the project structure in general. However, on the 'AI for Air interface' we prefer an initial study phase to align the details of use cases and assumptions on AI deployments. Additionally it is beneficial to have an independent SI on AI for NG-RAN to study the use cases not covered by Rel-17 SI.

2.2 Use Cases for AI/ML for Air-Interface SI

A large number of use cases were brought up. The most popular ones are summarized as follows:

- CSI feedback compression (lower overhead)
- Beam management (beam selection, beam recovery...)
- Positioning
- RS overhead reduction (channel estimation)
- Mobility
- Channel prediction

Some other use cases were brought up but received considerably less support including:

- Link Adaptation
- Power savings
- Receiver enhancements
- Scheduling enhancements
- Non-linearity tracking
- Auto encoders
- App-specific PHY optimizations
- MU-MIMO

The moderator proposal is to further discuss the following use cases:

- CSI feedback compression
- Beam management
- Positioning
- RS overhead reduction

Companies are welcomed to provide comments on the moderator proposal, as well as, to help refining what each one entails.

Feedback Form 8: Comments on Moderator's proposed use cases to focus on

1 – Futurewei Futurewei supports including the following use cases in Rel-18 SI: · CSI feedback compression · Beam management · RS overhead reduction
2 – vivo Mobile Communication Co. Having an initial set is ok if majority prefers this. On the other hand, we also believe use cases like mobility and CSI prediction useful. If we have an initial phase for use case study, it is possible to include more areas.
3 – AT&T Suggest not to downscope use cases at this stage. Certainly there is merit to the use cases listed by the moderator, however similar to the RAN3 study, some time at the beginning on the study item should be spent collecting and categorizing use cases as there may be commonalities in the approaches and requirements across them which would allow for a more general and inclusive study.
4 – vivo Mobile Communication Co. Similar view as AT&T above. One additional comment is that current selection of area is based on simple counting of companies' preference. More thorough discussion/study is needed for such selection.
5 – Convida Wireless We prefer the following use cases for Rel-18: CSI feedback compression Beam management RS overhead reduction Channel prediction We are open for other use cases as well.

6 – Qualcomm Incorporated

Qualcomm agrees with focusing on the four selected use cases for Rel-18 SI.

Regarding some proposals on having an initial pre-study phase to look into use cases, we prefer 3gpp agreeing on a limited number of use cases prior to the start of R18 SI. This already seems quite feasible given that companies already have good alignment on top use cases.

7 – Apple Hungary Kft.

Agree with the 4 use cases to be included in the initial SI scope. Further down-selection during SI is possible.

8 – CATT

In the initial phase of study, the listed use cases can be studied to gain better understanding of AI/ML and formulate proper scope of further study.

After the initial study phase, we suggest focus on limited number of use cases, for example, CSI feedback compression and beam management improvement can be included. Use case of RS overhead reduction is mainly related to receiver algorithm which is likely to be an implementation issue. Detailed study on this use case is not very helpful for the understanding of AL/ML for physical layer. Positioning is also a promising use case of AI/ML. But it is very different from CSI feedback compression and beam management. Including positioning use case would make the scope of the study too broad and involve a large group of experts on both positioning and MIMO.

9 – Intel Technology India Pvt Ltd

Support the moderator proposal on 4 use-cases

10 – Beijing Xiaomi Mobile Software

The 4 use cases are mainly about physical layer. We would like to study AI utilization in mobility management, which is mainly done by RAN2. Legacy mobility management may suffer from frequent failures or signaling overhead in dense network deployment, especially in FR2. AI based mobility enhancement could improve the robustness and save signaling overhead, which is very important for user experience.

11 – Intel Technology India Pvt Ltd

Un-precoded CSI-RS based rank and precoder feedback. Focus on large Tx ports. Spatial, frequency and time-domain compression can be considered. Rel-16 Type-II as a baseline.

12 – Beijing Xiaomi Mobile Software

Furthermore, we also would like to study AI based power saving mechanism. So, in general, we suggest not to do down selection in initial phase. The 4 use cases are too limited.

13 – Deutsche Telekom AG

Similar to other companies we would prefer to have an initial phase in the SI to select suitable use cases for AI/ML usage based on technical arguments (applicability of AI/ML models, possible benefits, ...). For the WID we can start a down-selection process to finally have a set of 6 to 8 use cases (out of the list provided by the moderator).

Focusing on only the 4 use cases proposed by the moderator at this point in time is too early.

14 – Spreadtrum Communications

We agree that the use cases listed by moderator are the most interested use cases among companies. But we share the same view with other companies, it is better not to down-selection use cases at current stage. We should set up one common basis for all use cases as possibly in Rel-18, e.g., data set construction, AI model, evaluation methodology and so on. We can down-select use cases during SI stage.

15 – Verizon UK Ltd

Agree with AT&T and DT views. Initial phase of SI can be devoted to use case selection based on technical merits and expected gains. Down selection for WI phase can be done at the end of the SI phase.

16 – NTT DOCOMO INC.

We support low overhead CSI feedback and beam management. We are not sure if companies are aligned with the meaning of CSI feedback compression. Our preference is not to apply AI/ML compression scheme of CSI feedback. Instead, AI/ML is deployed only on NW to reconstruct CSI with sparse CSI feedback as an input.

17 – LG Electronics Inc.

We wonder whether the moderator's intention is to narrow-down use cases before SI initiation. We think that AI/ML assumption and categories should be clarified first (e.g. 4 Cats provided by vivo or moderator's last proposal) before narrowing down use cases. Even for a single use case, specification impact would be quite different per this assumption/category. In this regard, we prefer to study general AI/ML framework in Rel-18 including assumptions/categorizations/use cases without limiting to a few use cases as suggested by AT&T, Vivo, and others.

18 – SHARP Corporation

We are fine with moderator's proposed use cases. We also agree that discussion and study on categorization of use cases and network/UE involvement are necessary, and it is not the time to preclude some use cases, considering the big picture. Meanwhile, we believe standardization of AI/ML for air interface is long-term work, which may be pursued for several releases. Thus, during Rel-18, it may be better to focus on constrained number of typical use cases, which could provide useful knowledge and methodology for the subsequent standardization work in the future releases.

19 – CAICT

We support at least the 4 use cases for further discussion. We also observe that some other use case, like channel prediction, can be combined with other use cases for better performance. It might be helpful to have some constraint on the study directions for further use case selection. But it is still recommended to have a use case study phase for further use case down-selection for performance evaluation.

20 – Guangdong OPPO Mobile Telecom.

We support study on CSI feedback compression, Positioning and RS overhead reduction. Beam management is too complicated to be improved by AI/ML at this stage in our opinion after initial study.

21 – Fujitsu Limited

Similar view as AT&T, it seems too early to decide use case candidates right now. We suggest open the door to other use cases during initial study phase. Down selection can be done after (or in parallel to) the study/conclusion of high-level principles in AI for air interface among companies.

22 – Fraunhofer IIS

We support to study Positioning and Beam Management. We are open for other use cases as well (lower priority)

23 – ZTE Corporation

We are supportive to focus on these four use cases now. We also support to include these four areas in the Rel-18 SI. Details for each use case can be further discussed and specifically described in the SID.

24 – MediaTek Inc.

MediaTek support to only focus on selective use cases in Rel-18.

From workshop input from companies, AI-assisted CSI feedback and AI-assisted mobility received the most support, so we support to start our study with these two.

25 – NEC Corporation

We supports the use cases in the moderator's proposal. We also propose that mobility could be included in the use cases.

26 – KDDI Corporation

Agree with AT&T, DT and Verizon views. It would be better that in SI performance gains of each use cases are firstly studied before down selection for WI phase.

27 – Sony Corporation

The use case list is OK to us. However, in line with views expressed by other companies, we think use cases shouldn't be down-selected at this stage. We should have an initial phase of the study to identify use cases.

28 – Nokia Corporation

While the workload should be kept reasonable, we do not see that the best way of doing so is to narrow down use cases already in RAN level now before any studies are conducted. We do not know how well these use cases represent challenges and opportunities of the AI/ML for air interface and thus, how well these particular use case proposals allow to study means of enabling feature-independent AI/ML signalling and workflow. So far there has not been any real technical discussion on use cases (e.g., companies do not yet have. a common understanding of what each use case means and implies – e.g. CSI reporting) and which use cases make sense for this study and for overall understanding of AI/ML for air interface. In our view the use cases should be decided based on the technical studies.

29 – RadiSys

We support the 4 Use cases proposed by Moderator. We would like to include Mobility use case for the initial set.

30 – China Mobile Com. Corporation

We suggest to select one or two use cases from the listed popular use cases. Based on the selected use cases, we can study the KPIs and evaluation methodologies. Our highest priority is CSI feedback compression.

31 – Lenovo (Beijing) Ltd

We suggest focusing on two use cases, at least in the beginning, in Rel-18 SI for air interface enhancement:

- CSI feedback compression
- Beam management

Though only two use cases are suggested as the starting point for evaluation, the general behavior should not be ignored, such as the UE and Network involvement issues mentioned in Section 2.4.

32 – HUAWEI TECHNOLOGIES Co. Ltd.

Ok to start from the four use cases listed above as high-level description, with possible further refinement to be done as part of the study. See below comments for each.

In addition, we are fine to have an initial phase in the SI to select suitable use cases for AL/ML usage also as suggested by some other companies, to ensure the selected use cases are the ones where most gains from AI/ML over existing implementations/mechanisms, though it is expected the 4 use cases proposed by the moderator are the most promising ones.

33 – Samsung Electronics Polska

In general, we think the use cases can be further study and identified in SI phase. It should allow companies to contribute and explain more details in SI, other than identifying the use cases during SI drafting phase without detailed studied. On the other hand, we are open to list some use cases as examples in SID or justification parts.

34 – Ericsson LM

We are fine with a general study of CSI and positioning and where RAN1 performs a deeper study on CSI, e.g. using evaluations and comparisons with baseline. The purpose is to learn how to address/evaluate future AI/ML feature proposals in RAN1.

We don't see a clear need for specification work on AI/ML based beam management and we don't see any major issue with performance of BM that AI/ML may address.

The RS overhead reduction seem to be of minor relevance, and whether there are any lessons learned from such study that are generally applicable to elevate RAN1 understanding of standardization of AI/ML is not clear to us.

Hence, we suggest to keep only "CSI feedback" and "Positioning" for further discussion.

35 – VODAFONE Group Plc

OK with the moderator proposal to study

- CSI feedback compression
- Beam management
- positioning
- RS overhead reduction

Feedback Form 9: Refinement on CSI feedback compression potential AI/ML Air Interface use case

1 – Futurewei

Futurewei supports using AI/ML-based approach to reduce CSI feedback overhead on Rx side and improve CSI reconstruction accuracy at Tx. The accuracy can be measured based on agreed-upon methodology and KPIs (to be discussed across companies).

2 – Intel Technology India Pvt Ltd

Un-precoded CSI-RS based rank and precoder feedback. Focus on large Tx ports. Spatial, frequency and time-domain compression can be considered. Rel-16 Type-II as a baseline.

3 – vivo Mobile Communication Co.

For detailed discussion in this area, at least the following aspects should be discussed:

- Sub-use cases in AI based CSI feedback compression use case, e.g., regarding level of coordination between NW and UE;
- Down-selection of sub-use cases for Rel-18 further study;
- Evaluation of the identified sub-use cases, including detailed evaluation assumptions, metrics for comparison, how to align DNN models between different companies and how to construct data set for training/validation/test;
- For each of the sub-use case, study framework and procedures on exchange of AI related information between network and UE that facilitate learning and inference;

The following are example sub-use cases for CSI feedback compression:

- Cat1 use case for CSI feedback compression: AI/ML is transparent to both sides. The network can conduct CSI prediction with sparser CSI feedback in time/frequency domain; Or the UE can use the AI/ML receivers to search for PMI;
- Cat2 use case for CSI feedback compression: AI/ML training is facilitated in spec non-transparent way, but the newly defined measurements or reporting behavior is not directly related to DNN information; For example, more CSI reports are collected to facilitate training at the network; Or more CSI-RS transmission is required for UE to train the AI/ML based PMI estimation;
- Cat3 use case for CSI feedback compression: Assistance information related to DNN is exchanged between two sides; For example, the network provides some DNN model that could expedite the application of the AI/ML based PMI estimation at the UE side.
- Cat4 use case for CSI feedback compression: Encoder-decoder like models are jointly used at the UE side and network side to compress the CSI feedback.

4 – NTT DOCOMO INC.

We think introducing collaborated AI/ML models deployed on NW and UE requires large specification impacts, such as model transfer. It is better to deploy ML models only on NW in Rel-18 as a first step of AI/ML. The coordination of ML models on both NW and UE, where UE applies compression scheme and NW reconstructs CSI with collaborated ML models, should be studied after Rel-18. In this sense, even though we support low overhead CSI feedback, compression scheme is not needed. Instead, we prefer

reconstructing CSI on NW based on sparse CSI feedback as a use case of low overhead CSI feedback. In this case, expected specification impacts are relatively low thanks to AI/ ML deployment on NW and exploiting existing codebook CSI feedback.

5 – Guangdong OPPO Mobile Telecom.

For AI based CSI feedback, the original CSI information can be compressed by an AI encoder located in UE, and recovered by an AI decoder located in gNB. KPIs, such as MSE(or Cosine Similarity) between the original information and recovered one can be used to evaluate the performance gain. In addition, BLER or throughput obtained from link-level simulations can also be used to evaluate the performance gain. Type 1 codebook and Type 2 codebook defined in Pre-Rel18 can be used as the comparative baseline.

Channels used for link level simulations and system level simulations(e.g. defined in TR 38.901) can be used to construct the train set, validation set and test set in this case, and need to be aligned among different companies

As proposed by many companies, complexity of the AI base approaches need to be evaluated in this case. The maximum computing complexity (in unit of [FLOPs]) needs to be limited, by considering the required inference latency under a reference computation capability (in unit of [FLOPS or TOPS]). The maximum model size needs to be limited as well considering the affordable storage complexity.

6 – CAICT

Full channel information(F-CSI) or partial CSI, like PMI could be compressed by AI/ML model at UE side and decompressed by AI/ML model at gNB side with limited overhead. F-SCI and partial channel information feedback are used as competition tasks in 1st WAIC and 2nd WAIC respectively. We are open to see more use cases for CSI feedback related use case with less specification impacts.

7 – ZTE Corporation

For CSI compression feedback, we support to study channel matrix or precoding matrix compression feedback in at least spatial domain and frequency domain. Rel-16 eType II can be used as the baseline scheme for the study.

- Study channel matrix or precoding matrix compression feedback in at least spatial domain and frequency domain.

  - Use Rel-16 eType II codebook as the baseline scheme

8 – Nokia Corporation

The same answer as for the Feedback Form 8.

9 – China Mobile Com. Corporation

We think we can study AI/ML based PMI-like CSI feedback, and compare the performance with TypeI/(enhanced)TypeII codebook based CSI feedback. The performance comparison should be done from the throughput perspective, not just accuracy.

10 – Lenovo (Beijing) Ltd

In this use case, we prefer to use AI/ML-based approach to compress the CSI with reduced feedback overhead and with reasonable loss on the reconstruction accuracy. The overhead reduction, loss on the accuracy and PDSCH performance based on the CSI can be defined as the agreed KPIs.

11 – HUAWEI TECHNOLOGIES Co. Ltd.

CSI feedback compression can be studied so that the signaling overhead of CSI feedback can be reduced or the accuracy of CSI recovery can be improved. It is preferred to use "CSI feedback with reduced overhead and/or improved accuracy".

12 – Samsung Electronics Polska

For CSI feedback compression, it is to compress the full CSI on Rx and reconstruct the CSI at Tx. Spatial-delay or spatial-frequency CSI can be studied as a starting point, comparing with Type-2 precoding. The AI/ML algorithm and model is supposed to be exchanged between UE and BS, dominated by BS.

13 – Rakuten Mobile

Agree with Moderators selection

- CSI feedback compression (lower overhead)
- Beam management (beam selection, beam recovery...)
- Positioning
- RS overhead reduction (channel estimation)
- Mobility
- Channel prediction

Would also like to Link Adaptation as a use case.

14 – Ericsson LM

Ericsson is supportive of an *AI-enhanced CSI reporting* use case focused on the following sub-topics (all with respect to Rel 16 and/or 17 CSI reporting as a baseline solution):

- Higher-resolution CSI reporting to enable improved DL spectral efficiency.
- Lower-complexity CSI reporting for UEs.
- More robust CSI reporting for high-mobility scenarios.

We are supportive of this use case for the following reasons:

- The use case has clear specification impact and (potential) performance improvements.

• The complexity of CSI reporting has long been an issue for UEs, and AI/ML technologies have the potential to reduce this complexity.

• Higher-resolution CSI reports, with comparable UL overhead, enable us to better exploit massive MIMO antenna arrays.

- The use case will likely require non-trivial AI/ML models in the UE and Network to interact with one another, and, therefore, will help "excite" learnings on AI/ML support, provisioning, and maintenance on the air interface.

Feedback Form 10: Refinement on Beam Management potential AI/ML Air Interface use case

1 – Futurewei

Futurewei supports using AI/ML-based approach to estimate/predict the best beam across all beam pairs that maximize the channel capacity.

2 – Intel Technology India Pvt Ltd

Fast beam-pair acquisition on downlink for single TRP operation

3 – vivo Mobile Communication Co.

For detailed discussion in this area, at least the following aspects should be discussed:

- Sub-use cases in AI based beam management use case, e.g., regarding level of coordination between NW and UE;
- Down-selection of sub-use cases for Rel-18 further study;
- Evaluation of the identified sub-use cases, including detailed evaluation assumptions, metrics for comparison, how to align DNN models between different companies, how to construct data set for training/validation/test and baseline for comparison;
- For each of the sub-use case, study framework and procedures on exchange of AI related information between network and UE that facilitate learning and inference;

The following are example sub-use cases for beam management:

- Cat1 use case for beam management: AI/ML is transparent to both sides. The network can conduct beam prediction with information reported with legacy procedure; Or the UE can use the AI/ML based beam search to reduce power consumption for beam management effort;
- Cat2 use case for beam management: AI/ML training is facilitated in spec non-transparent way, but the newly defined measurements or reporting behavior is not directly related to DNN information; For example, more beam measurement is required for training at the gNB.
- Cat3 use case for beam management: Additional information is exchanged between gNB and UE to expedite the training at the UE. For example, some models can be downloaded from gNB to UE as a reference for UE to train AI/ML models for beam search.
- Cat4 use case for beam management: Transmission of CSI-RS for beam management is compressed through neural network and reception of the CSI-RS is also based on neural network. The two neural networks are jointly designed and optimized.

4 – Spreadtrum Communications

For this use case, we could firstly focus on specific scenarios, e.g, train., where we think the benefit of AI/ML would happen with the highest probability. If the benefit identified to exist in these special scenarios, then we could consider to extend it to general scenario.

5 – NTT DOCOMO INC.

We support beam management with AI/ML. Enhancements of AI-aided beam management can be seen from two aspects: how to collect the optimal inputs for AI-aided beam management (e.g., enhanced beam measurement report) and how to exploit inference of AI/ML models (e.g., predicted beam indication and predicted BFR). The latter is under the assumption that AI/ML can estimate or predict accurate beam qualities.

6 – CAICT

AI/ML model could be used for completing the evaluation of all beam pairs according to the limited measurement beam pair information at UE side and choose the best beam(s) for further operation.

7 – ZTE Corporation

We think for beam management, beam prediction in time or spatial domain can be considered. The goal of this study is to improve UE experience in FR2 with reduced the beam measurement RS overhead and UE measurement/reporting.

- Study beam prediction in time or spatial domain to improve UE experience in FR2 with reduced the beam measurement RS overhead and UE measurement/reporting.

  - For beam prediction in time domain, focus on high mobility scenario (e.g., high-speed train and highway)

  - Study aspects include evaluate the feasibility and potential system level gain and specification enhancements such as beam measurement, beam report, beam indication and fall-back approach.

8 – Nokia Corporation

The same answer as for the Feedback Form 8.

9 – Lenovo (Beijing) Ltd

In this use case, we prefer to use AI/ML-based approach to determine and track/predict the optimal beam from the candidate beams based on measurements of the relevant reference signals. The overhead and latency reduction in comparison to existing framework can be selected as one of the KPIs.

10 – MediaTek Inc.

We support AI-assisted mobility, besides BM, we also supports using AI/ML-based approach to estimate/predict the best cell for for HO robustness and latency reduction.

11 – HUAWEI TECHNOLOGIES Co. Ltd.

Ok to include beam management starting from the current high-level description and open to further refine it, as it currently can include many sub-cases like beam training, beam tracking etc.

12 – Samsung Electronics Polska

For beam management, it includes AI/ML-based scheme to fast selecting, tracking and recovery Tx and Rx beam at both side for higher link quality. The scheme may be jointly or separately designed at BS and UE.

13 – Rakuten Mobile

AI/ML Based mobility and selection of Best Cell/Beam.

14 – Ericsson LM

We don't see a large need for further beam management enhancements in NR and therefore not the need to study standardized BM enhancements driven by AI/ML. Proprietary methods for AI/ML based beam management can of course always be implemented based on current specifications.

15 – RadiSys

We support AI/ML models for beam switching and best beam selection.

Feedback Form 11: Refinement on Positioning potential AI/ML Air Interface use case

1 – vivo Mobile Communication Co.

At least the following aspects should be discussed in this area:

- Sub-use cases in AI based positioning use case, e.g., regarding level of coordination between NW and UE;
- Down-selection of sub-use cases for Rel-18 focus;
- Evaluation of the identified sub-use cases, including detailed evaluation assumptions, metrics for comparison, how to align DNN models between different companies□how to construct data set for training/validation/test;
- For each of the sub-usecase, study framework and procedures on exchange of AI related information between network and UE that facilitate learning and inference;

For the AI based positioning, we think there are at least the following sub-use cases:

- Cat1 use case for positioning: AI/ML based estimate at the UE with training and inference transparent to the NW; Or AI/ML based TDOA estimate at the NW with training and inference transparent to the UE;
- Cat2 use case for positioning: Additional assistance information is exchanged between gNB and UE for AI based positioning. The information is not directly related to neural network, for example: gNB antenna information/calibration information can be provided to UE to improve positioning estimate accuracy based on AL/ML;
- Cat3 use case for positioning: Assistance information directly related to neural network is exchanged between gNB and UE. For example, trained/aggregated at network for the environment are distributed to UE to expedite the training at the UE; UE uses a neural network for positioning related estimate with the neural network derived from the assistance information.
- Cat 4 use case for positioning: Joint inference at gNB and UE. For example, UE side can extract features and reported to network that would further use by NW for inference.

2 – Intel Technology India Pvt Ltd

- LOS/NLOS classification and multi-path mitigation for DL positioning.
- Study assuming current architecture (measurements at UE reported to LMF)

3 – Spreadtrum Communications

In Rel-18, we suggest to focus AI/ML training and inference at gNB side, not UE side.

4 – Guangdong OPPO Mobile Telecom.

For AI based positioning, the obtained measurement results and/or channel states (e.g. RSRP/RSSI, or CIR/PDP/CTF) can be used for an AI based positioning model. Both UL based positioning and UE assisted DL based positioning need to be considered in this case. KPIs, such as accuracy achieved @[90]% as defined in TR 38.857, can be used to evaluate the performance gain. traditional approaches that used in Rel-16/17 can be used as the comparative baseline.

Scenarios that defined in 3GPP TR 38.857 can be reused as the AI based positioning scenarios. NLoS scenarios need to be evaluated first, since in Rel-17, it is clearly that to meet the 3GPP requirements on NLoS positioning through traditional approaches is a very challenging task.

As CSI feedback compression case the maximum computing complexityneeds to be limited The maximum model size needs to be limited as well

5 – CAICT

In general, AI/ML model could be used for positioning in LOS/NLOS scenario with channel impulse response as input to improve positioning accuracy. The detail of AI/ML model location and assistant information transfer could be FFS.

6 – ZTE Corporation

We support to include study positioning accuracy enhancement using AI/ML, especially in the scenarios with high NLOS probability.

- Study the improvement of positioning accuracy enhancement in scenarios with high NLOS probability, e.g., IIOT DH

7 – Nokia Corporation

The same answer as for the Feedback Form 8.

8 – HUAWEI TECHNOLOGIES Co. Ltd.

Ok to include Positioning starting from high-level description and open to further refine it. Both LOS/N-LOS identification, positioning performance for NLOS scenarios can be interesting topics.

9 – Ericsson LM

The positioning based studies can be included in a general study to asses methodologies, data sets etc,. No need for a deeper study of positioning using evaluations and performance assessment.

Feedback Form 12: Refinement on RS overhead reduction potential AI/ML Air Interface use case

1 – Futurewei

Futurewei supports using AI/ML-based approach to perform CSI-RS compression at Tx side and channel estimation (to recover the full channel matrix) at Rx side.

2 – vivo Mobile Communication Co.

For the involved RS type, we see the potential for DMRS, CSI-RS and SRS.

For detailed discussion in this area, at least the following aspects should be discussed:

- Sub-use cases in AI based RS overhead reduction use case, e.g., regarding level of coordination between NW and UE;
- Down-selection of sub-use cases for Rel-18 further study;
- Evaluation of the identified sub-use cases, including detailed evaluation assumptions, metrics for comparison, how to align DNN models between different companies and how to construct data set for training/validation/test;
- For each of the sub-use case, study framework and procedures on exchange of AI related information between network and UE that facilitate learning and inference;

The following are example sub-use cases for DMRS:

- Cat1 use case for positioning: AI/ML based channel estimation at UE side is transparent to NW; Or AI/ML based channel estimate at the NW is transparent to the UE;
- Cat2 use case for positioning: New DMRS patterns are designed to exploit the gains provided by AI/ML. E.g., DMRS with lower overhead in time domain or frequency domain;
- Cat3 use case for positioning: Assistance information directly related to neural network is exchanged between gNB and UE. For example, neural networks are trained/aggregated adaptively to the environment by the NW, and the information is transferred to UE to expedite the training at the UE; UE uses a neural network for DMRS channel estimate with the neural network derived from the assistance information.
- Cat 4 use case for positioning: joint inference at gNB and UE. For example, DMRS compressed through neural network at gNB side and de-compressed with the corresponding neural network at UE side.

3 – vivo Mobile Communication Co.

Some typos in previous comment is corrected.

For the involved RS type, we see the potential for DMRS, CSI-RS and SRS.

For detailed discussion in this area, at least the following aspects should be discussed:

- Sub-use cases in AI based RS overhead reduction use case, e.g., regarding level of coordination between NW and UE;
- Down-selection of sub-use cases for Rel-18 further study;
- Evaluation of the identified sub-use cases, including detailed evaluation assumptions, metrics for comparison, how to align DNN models between different companies and how to construct data set for training/validation/test;
- For each of the sub-use case, study framework and procedures on exchange of AI related information between network and UE that facilitate learning and inference;

The following are example sub-use cases for DMRS overhead reduction:

- Cat1 use case for DMRS overhead: AI/ML based channel estimation at UE side is transparent to NW; Or AI/ML based channel estimate at the NW is transparent to the UE;

- Cat2 use case for DMRS overhead: New DMRS patterns are designed to exploit the gains provided by AI/ML. E.g., DMRS with lower overhead in time domain or frequency domain;
- Cat3 use case for DMRS overhead reduction: Assistance information directly related to neural network is exchanged between gNB and UE. For example, neural networks are trained/aggregated adaptively to the environment by the NW, and the information is transferred to UE to expedite the training at the UE; UE uses a neural network for DMRS channel estimate with the neural network derived from the assistance information.
- Cat 4 use case for DMRS overhead reduction: joint inference at gNB and UE. For example, DMRS compressed through neural network at gNB side and de-compressed with the corresponding neural network at UE side.

4 – Intel Technology India Pvt Ltd

Overhead reduction for CSI-RS for CSI

5 – Spreadtrum Communications

In order to avoid to cause impact to legacy UEs, we suggest to focus on the UE specific RS overhead reduction, e.g., DL/UL DMRS, SRS. The time and frequency density could be considered to be reduced.

6 – Guangdong OPPO Mobile Telecom.

For AI based RS reduction, the obtained RS(e.g. CSI-RS, DMRS, SRS) can be utilized by an AI model to estimate the wireless channel information. The number of RS resources that is needed by AI based solutions and that is needed by traditional solutions can be evaluated to justify the performance gain in this use case. For example, e.g. with the same RS overhead, AI based solution could obtain [xxx] performance gain in [MSE], or with the same performance gain, [yyy%] RS resources can be reduced by AI based solutions. In addition, BLER or throughput obtained from link-level simulations can also be used to evaluate this case.

Reference signals(e.g. CSI-RS, DMRS, SRS) and Channels used for link level simulations and system level simulations(e.g. defined in TR 38.901) can be used to construct the train set, validation set and test set and the corresponding label tag, and need to be aligned among different companies.

As CSI feedback compression case computing complexity of the AI based approaches need to be evaluated also in this case. The maximum model size needs to be limited as well

7 – CAICT

RS overhead reduction is a relatively big area. AI/ML model could be used for DMRS/CSI-RS/SRS channel estimation. We observe there are two directions for AI/ML based channel estimation. One direction is using AI/ML model to achieve higher channel estimation accuracy with legacy RS pattern, which is selected as one task in 2nd WAIC. Another direction is making channel estimation by AI/ML model with new RS pattern. We support the study of both directions.

8 – ZTE Corporation

We support to study at least DM-RS overhead reduction for lower density at least in frequency domain, based on AI assisted channel estimation. We are open to reduced density in time domain and other RS as well, e.g., CSI-RS, SRS, etc.

- Study RS overhead reduction for lower density based on AI assisted channel estimation.
  - At least include lower density in frequency domain for DM-RS

9 – Nokia Corporation The same answer as for the Feedback Form 8.
10 – HUAWEI TECHNOLOGIES Co. Ltd. We think this item should be changed to "RS enhancement". AI can be applied to either reduce the overhead or enhance the overall performance.
11 – Samsung Electronics Polska In our understanding, the study can include the study on using AI/ML-based scheme to reduce the DMRS overhead for receiver channel estimation, in order to recover the full channel information in context of MIMO transmission.
12 – Ericsson LM The RS overhead reduction use case seem to be of minor relevance, and whether there are any lessons learned from such study that are generally applicable to elevate RAN1 understanding of standardization of AI/ML is not clear to us. We don't see a strong need to include RS overhead reduction in the list of use cases.
13 – Rakuten Mobile We see potential for DMRS, CSI-RS and SRS reduction through AI/ML

2.3 Evaluation Methodology and KPIs for AI/ML for Air-Interface SI

From the received companies' views, there is pretty broad support for basing the evaluation methodology on existing 3GPP framework for evaluations, i.e., channel models, link and system level simulations, etc.

A number of companies brought up the importance to not only focus evaluations on synthesized channel models but to add field data to assess the performance in real-world environments.

The KPIs are broadly understood to be use case specific but a number of common metrics prevail, e.g., complexity and performance vs. proper (non-AI/ML) baseline.

As a result, the following high level proposals are put forth for further discussion:

- Base the evaluation methodology in AI/ML for Air-Interface study on existing 3GPP framework for evaluations, i.e., channel models (TR 38.901 and TR 38.857 [positioning]), link and system level simulations, etc.
- Additionally, use field data to assess the performance in real-world environments.
- KPIs are broadly understood to be use case specific but a number of common metrics prevail, e.g., complexity and performance vs. proper (non-AI/ML) baseline.

Companies are welcomed to provide inputs to the above high-level proposals on Evaluation Methodology and KPIs for AI/ML for Air-Interface study:

Feedback Form 13: Companies input on Moderator's Proposals for evaluation methodology and KPIs for AI/ML for Air-Interface SI

1 – Futurewei

· For datasets, Futurewei supports using simulation data based on 3GPP defined channel models and using field data together for performance evaluation.

· For KPIs, they may be specified on per use case groups according to the task that AI/ML-based solution is designed to solve/perform. In addition, AI/ML model complexity and communication overhead introduced should be considered as part of the evaluation criteria.

· Proper/common non-AI/ML based solution(s) should be identified as baseline on per use case basis.

2 – vivo Mobile Communication Co.

We are fine with the directions of basing the evaluation methodology in AI/ML for Air-Interface study on existing 3GPP framework for evaluations. Moreover, it should also be possible to adapt the channel model when constructing data set for different purposes, for example, the data set for verification may be constructed based on some perturbation of parameters to verify generalization performance.

For collecting the filed data in real-world, we have concerns on its feasibility. We don't know how this could be pursued, maybe the proponents could clarify a little bit. Also we don't know whether the shared real world filed data would be representative enough for evaluation purpose.

3 – Qualcomm Incorporated

Qualcomm agrees with the high level proposals.

We want to note that the use of synthetic dataset is mainly for convenience; care should be taken in interpreting the results, with the understanding that the results on synthetic dataset may not generalize and that the model complexity and performance may change significantly when faced with real world dataset. Therefore, it's important to also look at field data to assess the performance in real-world environments.

On the performance KPI, a "proper (non-AI/ML) baseline" should be non-trivial and competitive baseline in order to properly judge the real merit of AI/ML-based approaches.

4 – Apple Hungary Kft.

We agree with high-level proposals on evaluation methodology. We would like to clarification on using field data as data set. Is this per company's own choice of field data, or 3GPP will publish a field data set so all companies can evaluate?

On AI model complexity, MACC can be used to evaluate complexity.

5 – CATT

Simulation data based on 3GPP framework could be used as starting point of the study. But a proper mechanism of differentiating training data and testing data shall be developed.

6 – Intel Technology India Pvt Ltd

Suggest replacing "channel models" → "statistical channel models" from TR. Suggest to replace "field data to assess performance" by "field data to assess performance and robustness"

7 – Beijing Xiaomi Mobile Software

In general, we agree the existing 3GPP framework should be reused. However, we also see some difference in AI evaluation. Data set is fundamental for AI training and performance. Even with the same AI model/algorithm, the performance may be different depending on data set. In other AI areas, common data set is widely used to evaluate the AI performance. To make fair evaluation, we suggest to build a common data set.

8 – Deutsche Telekom AG

We generally agree with the proposed high-level principles for the evaluation methodology, but we see the need for further clarification on the usage of field data, e.g.:

- Who will provide that data?
- Will there be data sets for different evaluation scenarios? If yes, how many scenarios are to be considered?
- Which information is required to be contained in those data sets? AI/ML models from different proponents may rely on different input data.

9 – Verizon UK Ltd

We are fine with high level principles proposed here by the moderator.

10 – NTT DOCOMO INC.

□Data set: We support assessing the performance with field data on top of data generated from channel models.

□KPI: We agree that KPIs are broadly use case specific and some KPIs might be common over use cases.

11 – LG Electronics Inc.

We agree that KPIs are to be use case specific but we have questions similar with other companies on how to use/define field data.

12 – SHARP Corporation

We are generally fine with moderator's high level proposals. Data sets collected from real world could be helpful, at least for model validation/performance test, when data sets purely generated from standardized simulation models are used for AI/ML model training. With that said, how to collect the data properly (e.g., fair, common, and typical) is an issue to be discussed.

13 – Guangdong OPPO Mobile Telecom.

We are general fine with moderator's proposal apart from field data, for which we doubt whether it is feasible. If sufficient data need be collected in field, it may take too much time. Or it is not worthwhile for simulation. So to accelerate study, we think common reference data set derived from channel modelling by computer is more realistic approach.

14 – CAICT

We are generally fine with moderator's proposal, but with some concerns on using field data to assess the performance in real-world environments. In order to make a fair comparison between different companies, the construction of data sets should be consistent. How to use the field test data for performance evaluation in RAN1 needs further discussion.

15 – Fujitsu Limited

We are fine with using 3GPP channel models (38.901, e.g.) as the start point for AI/ML evaluation. How to test the AI's generalization capability should be studied and evaluated.

As to data from real world, it is good to have. But how many data from which kind of scenarios are enough to verify an AI model? We should study whether it is feasible first.

16 – ZTE Corporation

We are generally fine to moderator's proposal in high level. Some specific comments

- For data set, the simulation data based on existing 3GPP models like 38.901 or 38.857 should be mandatory to use, and field test data is optional to companies, as it is not easy for companies to acquire real-field data, and real-field data just shows performance in quite restrictive scenario.
- It is better to have some general principle in the SID, such as prioritizing offline training for evaluation, recommending parameter randomized channel for both SLS and LLS, etc.

17 – NEC Corporation

Moderator's proposal is appropriate for the consolidation discussion in September.

18 – Sony Corporation

We have concerns on the feasibility of using field data. Is there going to be a common database of field data on which AI models can be compared? Who would provide this data? Where would this data to be stored? How much data needs to be stored?

Since this will be a recurring issue for future study / work items involving AI/ML, how to provide and use field data in the evaluation of AI/ML could be one of the issues that is studied within the SI.

19 – Nokia Corporation

We agree to base evaluations on existing 3GPP SLS scenarios. At least some of the selected scenarios shall include time-variant effects (e.g. time-variant traffic, user mobility) as it needs to be evaluated how well AI/ML schemes perform under such real-life conditions. Assumptions for training/exploration shall be agreed as part of this SI. Incl. cases where training/exploration are conducted on a different setup than the one where used (ensuring some scalability). Stability, convergence, and scalability/generalization of AI/ML solutions are very important aspects, so the new KPIs for monitoring this should be included in the SI scope.

20 – China Mobile Com. Corporation

For data set, we think simulation data based on 3GPP channel model can be used, and a common data set for each use case can be provided.

For evaluation methodology, a typical model for calibration is necessary.

21 – Lenovo (Beijing) Ltd

- Re-using existing 3GPP evaluation frameworks such as channel model can be a good starting point, , to do comparison with the baseline as a first step. Subsequently, we can consider using field data for further assessment.

- For KPIs, they can be discussed and agreed once the use case is decided. Furthermore, the signaling cost to enable AI/ML function, such as the model transfer, training overhead, complexity to do inference, should be quantized and comparable with the baseline. Additionally, the required latency for on-line training (if needed) should also be considered per selected use case.

22 – MediaTek Inc.

We support the high level proposals.

23 – HUAWEI TECHNOLOGIES Co. Ltd.

Ok to start from the high level proposals step by step. The baseline and KPIs are general and open enough, while the "use" of field data could be an interesting direction subject to further discussion, as part of the study.

24 – Samsung Electronics Polska

We agree on the high level proposals.

At current stage, system level and link level simulator can be used for training and validation data generation, in existing framework, e.g. channel model of 38.901. For field dataset, it could be used for testing performance and generalization.

KPIs are task-specific. For each task, accuracy, overhead, and latency are generally applicable. Conventional rule-based algorithm in 3GPP framework can serve as the benchmark for each task.

25 – Ericsson LM

We are fine with the first bullet, as Ericsson supports the use of existing 3GPP frameworks for evaluations (e.g., use-case specific simulation scenarios and channel models). The exact nature of these simulations (e.g., link or system level) will depends on the use case, as usual. Ericsson also expects there will be a need to agree on parameter/configurations to generate appropriate datasets (e.g., training, development, and test datasets).

Regarding the second bullet, it needs to be revised to "whether and how to use field data needs further discussion" Ericsson is not supportive of "real world" field data in this study item. Our internal work indicates that synthetic models will be sufficient to evaluate the initial potential of AI technologies on the Air Interface and how they might be enabled via standardization. We think the SI should not focus on a "Kaggle-style competition" wherein ML models are finely tuned to specific datasets to extract maximum performance. Rather, we believe that the SI should help RAN1 form agreements on how to approach AI solutions and explore standardization issues (e.g., signaling, training), via an example use case study. Existing 3GPP evaluation frameworks should suffice for this purpose.

On the third bullet, the performance of AI-enhanced features should be compared to the relevant advanced baseline in Rel 15, 16 or 17 where state of the art Rel.17 UE is assumed for baseline comparisons (using maximized UE capabilities). The KPIs are use case specific.

26 – Rakuten Mobile

We agree with others that high level proposal is fine for now.

27 – VODAFONE Group Plc

Field data still seems useful in ensuring that the new concepts are likely to bring real benefits.

2.4 Further discussion on UE and Network involvement for AI/ML for Air-Interface SI

A number of good comments were received in the general comments section, as well as, in the corresponding UE and Network involvement section. It is becoming more and more obvious that there are different levels of information exchange enabling various levels of collaboration between nodes. Also, those different levels of collaboration will lead to various levels of air-interface specification impact. Thus, it is important to clearly understand them to further elaborate what the SI would look after.

As a result, the moderator believes that it would be good to establish the various levels of intern-node coordination and information exchange as follows:

0) No collaboration framework: AI/ML algorithms purely implementation based and not requiring air-interface changes/extensions.

1) Inter-node assistance to improve the respective nodes AI/ML algorithms. This would apply to UEs getting assistance from gNBs (for training, adaptation, etc.) and vice-versa. This level does not require model exchange between network nodes.

2) Joint ML operation between UEs and gNBs. This level requires AI/ML model instruction or exchange between network nodes.

3) Hybrid approach with pieces of 0), 1) and 2) above.

Companies are welcomed to share their views on different collaboration levels for further discussion. Also companies are welcomed to share their desired scope for the candidate Rel-18 project on AI for Air-Interface related to these various levels of collaboration.

Feedback Form 14: Companies views on categorization of inter-node collaboration levels

1 – Futurewei

For Rel-18 study, Futurewei suggests to first focus on use cases that have less interface impact while the discussion for different levels of collaboration can be carried across companies to reach better alignment.

2 – vivo Mobile Communication Co.

We have following two comments:

- There is one category missed, for example a new DMRS/CSI-RS/SRS pattern can be designed based on new AI-receiver. This AI-receiver is implementation based. Thus it does not belong to either level

0) or level 1).

- All the levels of coordination are possible and the benefits of different levels would be different for different use cases. It may not be possible to directly down-select at this stage.

3 – Qualcomm Incorporated

Qualcomm agrees with the general categorization of the various inter-node coordination and information exchange put forth by the moderator.

In general, for each use case and inter-node collaboration level, appropriate network and UE involvement should be discussed.

The "1) inter-node assistance to improve the respective nodes AI/ML algorithms. This would apply to UEs getting assistance from gNBs (for training, adaptation, etc.) and vice-versa" is one of the key aspects of the AI/ML air-interface and should be studied in Rel-18 SI.

The "2) joint ML operation between UEs and gNBs" is an essential element for the CSI feedback compression use case and should also be studied in Rel-18 SI.

To contain the scope of Rel-18 study, 3gpp can prioritize offline training for the purpose of performance KPI, while still considering online training for inter-node signaling and procedure studies.

The SI scope should include, in its RAN2/RAN3 component, defining general ML management procedures derived from the studied use cases.

4 – Apple Hungary Kft.

For type (0) use case, since there is no specification impact, it can be de-prioritized for SI.

For type (1) (2) and (3), we think it is better to clearly list the coordination is at training phase or at inferencing phase. For training, whether the assistance/co-ordination is used for online training, or offline training, and whether supervised or non-supervised or re-enforcement learning is applied. For example, with offline supervised learning, (1) (2) (3) can all be used for inferencing only.

5 – CATT

Cat 0 does not need air-interface changes. We don't have to consider this category in this study. If needed, a RAN4 study/work item would be suitable to define corresponding performance requirements.

If Cat 2 is included in this study, we should avoid joint training between gNB and UE which involves complicated information exchanges and leads to large system overhead.

6 – Intel Technology India Pvt Ltd

- 0) if no spec impact this option is not needed ?
- 3) the hybrid approach is unclear

7 – Beijing Xiaomi Mobile Software

In the first release, we suggest to focus on the methods with less interaction, i.e. option 0 and 1.

8 – Deutsche Telekom AG

Unclear what is meant by (0). AI/Ml models/algorithms are certainly out of scope of standardization, we should only focus on input/output data and, dependent on use cases, on AI/ML model handling/lifecycle management incl. possible info/model exchange between NG-RAN nodes or functions or between nodes and UEs. Therefore, we see especially (1) and (2) as most relevant scenarios. Details are use case dependent.

9 – Spreadtrum Communications

In our understanding, the leading task for Rel-18 is to identify intrinsic AI/ML could enable air interface and bring in benefit with reasonable price. All collaboration levels could be considered for the SID.

10 – NTT DOCOMO INC.

Thank you for clarifying the collaboration levels. Rel-18 is the advent of AI/ML models for air interface. As a first step, we think 0) and 1) should be prioritized in Rel-18 with considering the amount of specification impacts.

11 – Verizon UK Ltd

Agree with general categorization of approaches by moderator.

No need for Option 0 since it has no spec impact. Option 3 is not clear and might get complex for first pass of AI/ML for air interface. So we prefer to focus on Options 1 and 2. As Qualcomm pointed out SI should also include in its RAN2/3 part ML management procedures for the selected use cases and options.

12 – LG Electronics Inc.

We think that some categorization is really necessary and we are generally fine with the categories proposed by moderator, but we are somehow confused that we are in SI phase now. This categorization of AI/ML model needs to be a part of study as far as we understand.

13 – SHARP Corporation

We are fine with the categorization of inter-node collaboration levels. We consider such categorization is necessary and open to discuss on refinements. We believe that in Rel-18, it is better to focus on use cases of collaboration level that would require minor specification impact (but not zero specification impact, which should be not in the SI scope).

14 – Guangdong OPPO Mobile Telecom.

we think 0) is out of scope i.e. it is up to implementation. For 1) apart from training, inference information could be also exchanged over air interface. 2) we think AI/ML model instruction or exchange between UE and gNB is necessary instead of between network nodes since the study focus on air interface. If online delivery is necessary, then we need take SA1's study into account

15 – CAICT

We could have some level of categorization to have a better understanding of UE and network involvement. Considering the relationship of performance and categorization, it might not be very suitable to restrict use cases selection according to the proposed categorization level.

16 – Fujitsu Limited

We don't think it is necessary to have such kind of category right now. It can be discussed as STD impacts per use case.

17 – ZTE Corporation

We are generally fine with the categorization. Further refinement can be discussed in the SI. Different use cases may require different levels of UE and NW inter-operation. For example, beam management, RS overhead and positioning may require 0) or 1), and CSI feedback may require 2). Hence we don't think it is good to discuss selection of these categories now. All these categories should be kept for now. Once we have a concrete set of use cases for study in Rel-18, we can discuss further on how to select different levels of UE and NW inter-operation during the SI.

18 – NEC Corporation

To have a quick understanding of AI for the air interface and to facilitate 3GPP discussion, we prefer simple frame work in the first step. So we support 0 and 1 in the moderator's proposal. And we don't agree to rule out option 0 at SI phase, because if option 0 achieves similar performance as other options in the evaluation, the air interface changes in other options do not seem necessary.

19 – Sony Corporation

Discussion of the amount of inter-node collaboration should be one of the outputs of the study item. We should study the types of inter-node collaboration that are required (the list from the moderator is a good starting point) and then determine which use cases require which sort of inter-node collaboration.

20 – Nokia Corporation

It is good to study different levels of collaboration. Furthermore, as discussed already in the Rel-18 RAN WS, it is important to study in the early phase how to ensure and verify predictable UE behaviour with AI/ML and that the UE utilizing AI/ML for air interface meets all the existing UE minimum requirements while also obtains improvements in performance and minimum requirements with help of AI/ML.

We also have some clarifying questions and considerations. It is our understanding that 0) does not require any 3GPP involvement (or specification). Do others share this view? If yes, then how do we consider hybrid approach, which include some components of type 0) ? It should still be possible to define UE requirements for those parts, which are specified.

21 – China Mobile Com. Corporation

The levels of inter-node coordination and information exchange are different for each use case. Moreover, if model update or online training is needed, 1) and 2) should be prioritized. We also think this issue can be discussed after the use case(s) is determined.

22 – Lenovo (Beijing) Ltd

We are supportive to have a study on all possible collaboration levels, e.g., to support distributed AI.

For Level 2), it requires AI/ML model instruction or exchange between network nodes. As the answer in Feedback Form 7, we doubt this part can be handled in the scope of AI/ML for air interface. In some use cases, the AI/ML model management between network nodes might not involve air-interface at all. Thus, we prefer to have a separate SI, probably led by RAN3 or included in AI/ML for NG-RAN, focus on the

collaboration levels to better understand the merits/gains of the different collaboration frameworks for the identified use cases.

23 – HUAWEI TECHNOLOGIES Co. Ltd.

Our view is that it is too early to identify the inter-node collaboration levels at this stage. As we mentioned in the initial round, UE and Network involvement for AI/ML for Air-Interface should be discussed for each specific use case, by considering the procedures of model training and model inference. The signaling between network and UE can also be different for each use case. The corresponding details need to be ironed out during the study item. Nevertheless, our view is that detailed AI/ML algorithms and models are left for implementation.

24 – Samsung Electronics Polska

Whether collaboration between nodes is needed or not is depending on different use cases. It's not easy to evaluate without specific use case. This should be part of work in the study item.

25 – Ericsson LM

We are fine with this categorization with the following comments.

Category 0 (no collaboration framework): This appears to concern proprietary solutions operating within the existing standard. We do not believe this category needs to be included in SI. Reference baselines for the use case study can use the relevant Rel 15, 16 or 17 solutions where state of the art Rel.17 UE is assumed for baseline comparisons (using maximized UE capabilities). Hence, item should be modified to "0) No collaboration framework: AI/ML algorithms purely implementation based and not requiring air-interface changes/extensions. Hence, no standardization impact foreseen"

Categories 1, 2, and 3: Ericsson is supportive of such discussions and categorization. The example use case can be preceded by a high-level discussion on these categories (i.e., what is feasible).

26 – Rakuten Mobile

We agree with Oppo that 0) is out of scope, not sure about the role of 3GPP standardization.

Rest we agree with proposal.

27 – VODAFONE Group Plc

We tend to agree with Verizon.

3 Final Round

Many thanks for all the inputs received during the Intermediate round of discussions.

3.1 AI/ML Projects

Many companies have agreed with the proposed project structure in the Intermediate Round. A number of comments were received for the desire to further study new NG-RAN AI/ML uses cases in RAN3.

Some replies to some of the comments received:

- Moderator suggests not to further discuss the scope of the Rel-18 candidate RAN3-led **AI/ML for NG-RAN WI** because it should be based on the outcome of the on-going SI.
- Whether or not to agree with some **high level principles** for the candidate project on AI/ML for Air-Interface is to be discussed in this Final Round.
- Whether or not to have a **pre-Study phase** to discuss use cases for the candidate project on AI/ML for Air-Interface can be discussed when we get closer to the drafting of the corresponding SID objectives. Note that spending too much time further discussing use cases would limit the remaining time to study the true 3GPP implications to support AI/ML in the Air-Interface, while the ”preferred” use cases would likely be still the same as we have currently identified.

Based on these comments, the Moderator AI/ML project structure for discussion during the Final Round for discussion follows. Particularly, please indicate whether you agree with the addition of Project 3 below, as compared to not having it for the time being:

Table 2: AI/ML related candidate projects for Rel-18

	Temporary Title	**SI/WI**	**Primary WG**	**Secondary WGs**	**Notes**
1	AI/ML for NG-RAN	WI	RAN3	RAN2	Based on the outcome of RAN3-led Rel-17 SI *FS_NR_ENDC_data_collect*.
2	AI/ML for Air-Interface	SI	RAN1	RAN2, RAN4	SI for entire Rel-18 time-frame focusing on limited identified use cases, preferably one.
3	Additional Use Cases for AI/ML for NG-RAN	SI	RAN3		Study of additional use cases for RAN3 project.

Feedback Form 15: Companies inputs on revised project structure for AI/ML items in Rel-18

1 – Verizon UK Ltd

We agree with the overall set of activities listed by the moderator including addition of Project 3.

2 – Futurewei

Futurewei supports setting up a separate Rel-18 SI to study additional use cases for NG-RAN that are not covered under existing Rel-17 SI: FS_*NR*_ENDC_*data*_collect.

3 – Qualcomm Incorporated

We agree it is useful to continue the use case study. If the outcome of study is only on data collection and SON enhancement, we can consider to merge these into the R18 MDT/SON. If the use cases also involve air interface changes, we can also consider to merge with project 2.

To justify a new project in RAN3, based on the use cases study, we can study the general purpose procedures for model training, model inference and data management. For better inter-vendor interworking, flexible deployment and better efficiency, we can study the RAN architecture enhancement for RAN AI/ML, somewhat similar to what SA2 did for AI/ML in core network.

4 – China Mobile International Ltd

As Rapporteur of the SI: FS_*NR*_ENDC_*data*_collect, we share the same view as the moderator, suggest not to further discuss the scope of the Rel-18 candidate RAN3-led AI/ML for NG-RAN WI, and also not to further discuss the cotinuation of further study on AI for NG-RAN**,** because all the two parts should be based on the outcome of the on-going SI.

5 – Apple Hungary Kft.

Apple: Agree with the proposed structure

6 – LG Electronics Inc.

Support the proposed structure including the addition of one SI to cover the remaining NG-RAN use cases and the study of potential NG-RAN architecture enhancement for AI

7 – Intel Technology India Pvt Ltd

1) Agree.

2) generally OK and agree that further down-selection or refinement of use-cases can be done closer to SID preparation time (and minimize such efforts in WG level)

3) generally Ok. Besides new use cases, we are also ok to consider further architecture and framework enhancement in RAN3 SI.

8 – NTT DOCOMO INC.

We agree with all the projects (including project 3) listed by moderator.

9 – Guangdong OPPO Mobile Telecom.

we are fine with RAN3 WID, but we don't think it is necessary to spend more time on new use cases considering majority companies are more interested in RAN1 leading use cases. Therefore we don't think new study on additional use cases of RAN3 project is necessary. For RAN1 leading SI, we agree use cases should be limited, but it is too early to say "preferably one". So we prefer to remove the wording "preferably one"

10 – Deutsche Telekom AG

We are fine with the proposals for a SI on AI/ML for Air Interface and a WI on AI/ML for NG-RAN plus an additional SI on further use cases not considered within RAN3's Rel-17 study.

We would like to emphasize - similar to Qualcomm and Intel - that also enhancements on framework and architecture should be considered in RAN3's Rel-18 SI (see our inputs to former discussion rounds).

11 – ZTE Corporation

The split of three projects is fine for us. Comments on some details.

- On AI/ML for air interface, it's too early to decide the number of use cases now. Without the detailed scope of each use case, it impossible to assess how many use cases we can accommodate, although we agree we should be careful and just select a limited set of use cases. Hence we prefer to remove "preferably one".
- We support to discuss other high-layer use cases and other enhancements for AI in NG-RAN in a new SI led by RAN3.

12 – SHARP Corporation

We tend to prefer the table with two projects during intermediate round because we do not see strong justification on further dividing the projects.

13 – Fraunhofer IIS

We agree with the proposed structure by Moderator.

14 – Spreadtrum Communications

Support moderator's proposal

15 – CAICT

We are fine with the addition of Project 3.

We also would like to remove "preferably one" on AI/ML for air-interface

16 – Fujitsu Limited

We think whether to have Project 3 would be better based on the discussions and outcome of the on-going SI in RAN3.

17 – CATT

We support to have addition of project 3 which could be the continuation of the current Rel-17 SI.

18 – RadiSys

We agree to the project structure proposed by Moderator

19 – NEC Corporation

We support all three directions summarized by the moderator for AI/ML in Release 18.

20 – Ericsson LM

We support 1 and 2.

For 3, we can be supportive if the Study Item focuses on use cases and respects the high-level principles from the Rel-17 study item. It should have low priority.

21 – ROBERT BOSCH GmbH

We agree with the proposed structure by the moderator including the SI for the New Use Cases.

22 – vivo Mobile Communication Co.

We are fine with the additional project for more use cases.

For the project on air interface, whether we have ”pre-phase” or not is still subject to further discussion, thus it may not be appropriate to say ”SI for entire Rel-18 timeframe focusing on limited identified use cases, preferably one.”. It is still possible that we have relatively broader areas to look into during pre-phase. Down-selection could happen after the ”pre-phase”. It can be discussed whether we have a focused limited case after the ”pre-phase”.

23 – Nokia Corporation

We agree to have the following three projects; AI/ML for NG-RAN WI (RAN3, RAN2), Additional Use Cases for AI/ML for NG-RAN (RAN3) and AI/ML for Air-Interface SI (RAN1, RAN2, RAN4) for entire Rel-18 timeframe.

However, like requested large number of companies during intermediate round further study on use cases should be conducted before deciding the use cases.

24 – Lenovo (Beijing) Ltd

We are supportive to have a SI for NG-RAN, in addition to the SI for air interface. The potential scope of this SI includes:

- Study AI management procedures, AI/ML model distributing/update, data sharing with security and privacy etc. among RAN nodes to support distributed AI/ML;

- Study High layer related new use cases (e.g., AI/ML for slicing, QoE and etc.);

- Study integration and collaboration of OAM AI/ML, 5GC AI/ML, NG-RAN AI/ML.

25 – Samsung Electronics Polska

We are general fine with the proposed structure of AI/ML SI/WIs.

26 – VODAFONE Group Plc

Use Case 3 would be useful for Operator Specific Data capture, so we agree with the inclusion of this set

27 – MediaTek Inc.

We have item 3 to further study on new data collection for performance optimization with merit-based justification, and also on architecture enhancements for training data aggregation among network entities, distributing computing, and efficient model delivery.

28 – HUAWEI TECHNOLOGIES Co. Ltd.

Whether to add an additional SI for additional use cases for RAN3 as suggested by project 3 above would depend on the out-come of the on-going RAN3 SI, thus it is premature to add project 3 at this stage.

29 – InterDigital

We agree with the project structure, including the addition of project:3 – '*Additional Use Cases for AI/ML for NG-RAN'*

30 – Sony Corporation

We are ok with the revised project structure.

However, we don't agree with the wording of the note that the SI should focus on limited identified use cases. We should have an initial phase of the "AI / ML for Air Interface" study to identify use cases.

The moderator would like to request companies' input on agreeing on a set of high-level principles for the candidate SI on AI/ML for Air-Interface. As suggested, we could start from the RAN3 agreed principles (from the ongoing SI on *FS_NR_ENDC_data_collect*:

- Detailed AI/ML algorithms and models are left for implementation;
- User data privacy needs to be preserved;
- Support of AL/ML shall reuse the existing RAN architecture and new interfaces shall not be introduced.

Feedback Form 16: Companies inputs on high-level principles for the candidate SI on AI/ML for Air-Interface

1 – Verizon UK Ltd

Agree with proposals 1 and 2. While detailed AI/ML algorithms, models/training should be left to implementation to foster innovation, interface support to AI/ML model life cycle management as well as model training should not be precluded.

2 – Futurewei

In general, Futurewei is ok the high-level principles identified under FS_*NR*_ENDC_*data*_collect. Additional principles may be discussed during Rel-18 SI phase on AI/ML for Air-Interface.

3 – Qualcomm Incorporated

We agree with the first principle that detailed AI/ML algorithms and models are left for implementation. However, we don't see why we need to define this principle, as it's already clear that 3gpp won't specify any algorithms/models during a SI phase.

We agree with the second principle that user data privacy needs to be preserved. However, we do not see strong relevance of user privacy to the AI/ML for Air-Interface and therefore we do not see reason for having this principle. For example, how is compressed channel feedback or beam related feedback relevant to user privacy?

We do NOT agree with the third principle. One of the goals of this R18 SI is to identify the need of any changes in RAN architecture or the need of new interface. We can't conclude before studying.

In conclusion, we don't see need of defining the above principles.

4 – China Mobile International Ltd

We don't see the need to agree on these principles now, these should be the outcome of the study. It is still not clear whether these principles are needed for AI enabled air interface, specifically,

1. The first proposal is related to issue 3.4. More importantly, if the detailed AI/ML algorithms and models are fully left for implementation, how to ensure the performance in practical network? at least for evaluation and calibration, we need to touch somehow the AI model and algorithms.

2. We also don't see the strong relevance of user privacy with AI for air interface, e.g., CSI feedback compression.

3. The existing RAN architecture can be the baseline, but some enhancements should not be preclude, if neccesary.

5 – NTT DOCOMO INC.

We agree to reuse RAN3 agreed principles, as long as it fits selected use cases without any impediment.

6 – Apple Hungary Kft.

Agree with the principles.

7 – LG Electronics Inc.

Fine in principle but we are not sure how this can impact SI objective. It would be ok to discuss/decide this during SI phase as China mobile suggested.

8 – Intel Technology India Pvt Ltd

OK with understanding of bullet-1 that model architecture, dimensions, and inferencing may be discussed, training algorithms are left for implementation

9 – Guangdong OPPO Mobile Telecom.

we are general fine with 3 principles. We can further discuss during study whether new principle is needed

10 – ZTE Corporation

We are fine with the three high-level principles.

11 – Deutsche Telekom AG

Principles 1 and 2 are fine for us. With respect to (1) we should focus on input/output data for Model Inference and Model Training functions. (2) is probably of minor importance for AI/ML for Air Interface.

In view of (3) we have a different understanding. The existing RAN architecture should be used as baseline for Rel-18, but potential changes/extensions triggered by the needs of AI/ML model lifecycle management should not be excluded (part of RAN3's proposed Rel-18 SI).

12 – SHARP Corporation

We are fine with the listed high-level principles #1 and #2. Principle #3 requires futher check if no common understanding on the RAN architecture and new interfaces. Other necessary principles could be further discussed.

13 – Spreadtrum Communications

Agree in principle

14 – Xiaomi Communications

Xiaomi :

We think there is no need to limit the new use cases at the begining of the SI phase. The new use cases should be narrowed down based on fully study. And we don't understand why only one use case is perferable, it is too early to draw such conclusion. We suggest to remove the wording of "preferably one"

15 – CAICT

We are fine with moderator's proposals. In addition, we also think the unified network structures should be applied for both AI for air-interface and NG-RAN.

16 – Fujitsu Limited

We think it is necessary to have a set of high-level principles as the guidance for this study item. We may take the principle agreed in FS_*NR*_ENDC_*data*_collect as reference, and conclude the principles of AI/ML for Air-Interface.

BTW, the aggreged principles listed by moderator seems not include all from the latest RAN3 meeting?

17 – NEC Corporation

We agree to re-use RAN3 AI/ML principles as much as possible.

18 – Nokia Corporation

We agree generally with moderator's proposal on high level principles. However, in our view such details could be left open for the actual study item to define and work further. Furthermore, while we also agree that AI/ML algorithms and models are left for implementation, it is important that companies provide details for evaluating performance and complexity aspects for different proposals and approaches.

19 – Lenovo (Beijing) Ltd

We think the study objectives in RAN1 and RAN3 could be different for different scopes, so as for the principles.

For bullet 1, we agree that the AI/ML algorithms and models are implementation issues, but the study for air interface could have to do some evaluation on them in some degree of details. Thus, we need to define some basic principles to select the AI/ML algorithms/models and how details do we really need for comparison, which can be done in the evaluation methodology.

For bullet 2, it is a bit confusing on 'user data privacy' for a RAN1 study. For example, are the CSI/PMI and any other results regarded as the private data? We don't think this bullet make sense for the potential study areas for air interface.

For bullet 3, we can agree that in the initial phase (if we agree to have a phase-by-phase study) in the SI for air interface. If the benefits can be obtained from some new features for AI/ML operations, the new interfaces and architecture may not be excluded in the study.

20 – vivo Mobile Communication Co.

We are fine with the second principle.

For the first principle, our understanding is the model information exchange is one of the categories we need to study. It is not appropriate to preclude at this stage.

For the third principle, it is also pre-mature to conclude in this way.

21 – Samsung Electronics Polska

We agree with many other companies that, we can leave the discussion to SI phase. Different from RAN 3 led SI, the use cases and outcome of AI/ML for Air-interface might be quite different. For example, it may not related to user data privacy; if AL/ML algorithms need to be aligned at Tx and Rx, not sure whether it can be left for implementation; some new signaling/channel/signaling might also be needed for AI/ML for PHY. Therefore, we think SI is the right place to discuss the high level principles.

22 – VODAFONE Group Plc

In general we agree to the three principles

23 – MediaTek Inc.

We agree with the 3 principles.

24 – Ericsson LM

Ericsson supports the moderator's proposal

25 – HUAWEI TECHNOLOGIES Co. Ltd.

We agree with the three principles listed by Moderator and do think it should be captured to clarify the relationship between project 1 and project 2. In our understanding, in order to ensure consistent framework with compatibility and interoperability between features/projects, we do need to follow the same principle/framework, thus the ones agreed for RAN3 SI (i.e. project 1) should be applied for the SI of AI for Air Interface (i.e. project 2) also.

In addition, the following points are provided to clarify/response the comments from a few companies above:

- Principle 1 doesn't preclude setting up AI model for evaluation/calibration, the point is not to specify it in the specification. In addition, leaving it to implementation is actually to ensure the performance, while specifying detailed models in the specification will limit the potential performance that can be achieved.

26 – InterDigital

We support using RAN3 agreed principles as a starting point for the study on AI/ML for Air-interface.

27 – KDDI Corporation

Agree with proposal 1 and 2. On the other hand, proposal 3 seems to depend on use cases. If use cases need new interfaces, we support proposal 3.

3.2 Use Cases for AI/ML for Air-Interface SI

Companies expressed their views about the various use cases. A number of companies would like to do the down-selection of use cases as part of the SI while other companies are fine with a selection of a representative set of use cases ahead of the start of the SI, which is what the Moderator proposed.

Clearly the following use cases received the broadest support (beyond the point of whether individual companies see or not their favorite use case):

- CSI feedback compression (lower overhead)
- Beam management (beam selection, beam recovery...)
- Positioning
- RS overhead reduction (channel estimation)
- Mobility

All other use cases got considerably smaller support and the moderator fails to see the value of having an unconstrained list given the large number of use cases with large support. Indeed, the moderator agrees with some companies views that ideally we should identify a single use case to focus the study on.

Companies have provided additional information around the most popular uses cases which help better understand their meaning for other companies understanding.

At this point, the moderator thinks that use cases for a candidate Rel-18 SI on AI/ML for Air-Interface have been sufficiently debated during this email discussion and would like to try again the following proposal.

Proposal: Base further discussions on *use cases* for a candidate Rel-18 SI on AI/ML for Air-Interface on the following set of use cases:

- CSI feedback compression (lower overhead)
- Beam management (beam selection, beam recovery...)
- Positioning
- RS overhead reduction (channel estimation)
- Mobility

Feedback Form 17: Companies inputs on Use Cases for further discussion for the candidate SI on AI/ML for Air-Interface

1 – Futurewei Futurewei is ok to start from the proposed use cases as candidates for Rel-18 SI and down selection can be discussed as part of the SI. We suggest adding descriptive text for “mobility” use case in the AI for Air-Interface SI.
2 – CATT These use cases are used for discussion before Decemeber. The set of use cases included in the study after

December would depend on the output of the discussion. It that correct understanding?

3 – Qualcomm Incorporated

We agree to limit further discussions on use cases for "R18 SI on AI/ML for Air-Interface" to the above 5.

4 – China Mobile Com. Corporation

We support the proposal, and suggest to select one or two representative use cases ahead of the start of the SI.

5 – NTT DOCOMO INC.

All in all, we are fine with the proposal. However, we prefer replacing "CSI feedback compression" with "low overhead CSI feedback". "CSI feedback compression" implies deploying joint ML models on UE and NW, where UE applies compression scheme and NW reconstructs CSI with collaborated ML models. It is too early to confine only use cases with NW and UE strong involvements for AI/ML at this point. Applying AI/ML compression scheme of CSI feedback is not of our high interest. Instead, we prefer AI/ML deployed only on NW, where AI/ML reconstructs CSI with sparse legacy codebook-based CSI feedback as an input.

6 – Apple Hungary Kft.

Agree with the proposal. We support AI based mobility optimization to reduce handover failures and optimizing CHO, e.g. in terms of network resources reserved in candidate target cells.

In R17 RAN3 led SI, 3 prioritized use cases are being studied: energy saving, load balancing and mobility optimization. Further clarification of the different focus areas about this R18 RAN1 led SI on mobility use case versus R17 RAN3 led mobility use case is needed.

7 – LG Electronics Inc.

AI-RAN SI results cannot be directly applied to AI for air-interface so it is preferred to start from general AI framework for air-interface. If we really need to select use cases before SI, we are ok to start from the listed 5 use cases.

8 – Intel Technology India Pvt Ltd

1-4 is OK. With respect to mobility - we think RAN1 related work for mobility can be absorbed into beam-management and other mobility aspects can be part of RAN3 use-case

9 – Guangdong OPPO Mobile Telecom.

we can at least take CSI feedback compression and then down select the rest during study. In this way down selection of use cases and other work e.g. building the data set can proceed in parallel

10 – ZTE Corporation

We are fine to further discuss these 5 use cases. Detailed scope of the use cases subjects to further discussion and refinement.

11 – Deutsche Telekom AG

Ok to start wit that initial list of use cases in further discussions until Dec RAN Plenary. It is questionable if a down-selection to 1 or 2 use cases can be performed as this would require already a detailed analysis on impacts (e.g. AI/ML model handling only on NW side or spread across NW and UE, online/offline training, etc.).

What is missing is a short explanation of mentioned use cases. E.g. only stating ”mobility” or ”positioning” is certainly not sufficient.

12 – SHARP Corporation

We are fine with the use cases in the proposal in general. Our preference is to have a smaller set of use cases for Rel-18 SI, without precluding use cases that could be studied in the future.

13 – Spreadtrum Communications

We are fine with moderator’s proposal if it is the majority’s choice.

14 – CAICT

As a starting point of further use cases discussion, we agree to provide the list 5 use cases high priority. However, we are not clear if other use cases listed in intermediate round has a chance to be discussed somewhere. Some 0a use cases, like channel prediction, anyway, could be used at gNB/UE side to enhance scheduling/receiving performance.

15 – Fujitsu Limited

We think the use case ‘mobility’ should be further clarified, and whether it is suitable for a RAN1-led study item? We agree to have other 4 use cases as the start point for easy the discussion of evaluation methodology etc. But additional use cases should not be excluded before the start of the SI at least.

16 – Fraunhofer IIS

We are also fine with the identified use cases in the proposal. Similar to DT, a short explanation for some of the use cases is needed.

17 – NEC Corporation

We support these use cases for the SI.

18 – Lenovo (Beijing) Ltd

We agree with the above selected use case, but just a bit doubt they can be well evaluated in the SI with limited TU budget. Thus, we prefer to have at most two above use cases, i.e., CSI feedback compression and beam management, as the candidates in the SI for air interfaces, though there still need some alignment among companies on the AI/ML approaches for those.

19 – Nokia Corporation

In our view the decision of the use cases for the study should be part of the study item like requested number of companies during the intermediate round. More technical discussion is needed for deciding list of use cases.

20 – VODAFONE Group Plc

We are fine with the initial list, however we suggest another look at these 5 parameters at working group level for prioritisation.

21 – Samsung Electronics Polska

We strongly suggest to study and identify the use cases in SI phase other than limit it now. In SI phase, companies can bring more detailed analysis and results for each use cases. AI/ML is a new attempt for 3GPP, we should be more open for study and then down select some appropriate use cases to specify.

Moreover, current listed use cases are all heavy weighted-AI, which require offline training. The study item of AI/ML should also include another category, i.e., light weight-AI with online training. At least one use case on lightweight AI with online training should be included in this potential Rel-18 SI, e.g., non-liner handling.

22 – vivo Mobile Communication Co.

As we stated above, if we have a ”pre-Study phase”, broader areas can at least be considered during the ”pre-Study phase”. Down-selection can happen after the ”pre-Study phase” based on consensus from companies.

With the possibility of such ”pre-Study phase”, more use cases can be selected for discussion: besides above 5 use cases, another two use cases with high potential: CSI predication and PA non-linearity can be considered.

23 – MediaTek Inc.

Our suggestion on the use cases and corresponding WGs:

- CSI feedback compression (lower overhead) (R1)
- Predictive mobility, including Beam management (beam selection, beam recovery...) and handover (R2, R1)
- Positioning (R2)
- RS overhead reduction (channel estimation) (R1)

24 – Ericsson LM

We are fine to further discuss these identified use cases, with the understanding that further down-selection may apply before SID drafting.

25 – HUAWEI TECHNOLOGIES Co. Ltd.

1. We don’t see strong motivation to add mobility. Mobility is one of the use cases discussed in Rel-17 SI on AI for NG-RAN, it is not clear what additional thing we can do here.

2. We suggest to make the following changes to the use cases, since the current wording limits the scope for a certain use case. In our understanding, more careful study is needed in order to identify the most beneficial enhancements for a certain use case.

a) CSI feedback

b) Beam management

c) Positioning

d) RS enhancement

26 – Sony Corporation

We think use cases shouldn't be down-selected at this stage as other companies also mentioned in intermediate round. We should have an initial phase of the study to identify use cases.

An issue with considering a single use case is that that use case may not illustrate the full range of issues with supporting AI/ML in the air interface. We need to consider several use cases to work out what the potential impacts on the air interface are. For example, if the study just considered CSI operation with sparse CSI feedback, then the study could conclude that AI/ML could be implemented solely in the gNB without additional interaction from the UE. Such a conclusion would not be generally applicable, for example autoencoders may require significant signalling between the UE and gNB, even to the extent of model transfer.

In summary, the AI/ML study needs to consider a sufficient number of use cases to make some general conclusions on how AI/ML could impact the air interface.

27 – InterDigital

We think that proposed list of use cases can be considered as candidates for further down selection and preferably a single use case can be selected for the SI.

28 – ROBERT BOSCH GmbH

We support these use cases including mobility.

29 – KDDI Corporation

We agree with all listed use cases as a baseline for SI phase.

3.3 Evaluation Methodology and KPIs for AI/ML for Air-Interface SI

From the inputs received, there is a broad support to the Moderator's proposals in this section for the Intermediate round. Based on some of the comments received, the Moderator has refined the Proposal as follows (*italic* used for changes vs. Intermediate version):

Proposal:

- Base the evaluation methodology in AI/ML for Air-Interface study on **existing 3GPP framework for evaluations**, i.e., *statistical* channel models (*from* TR 38.901 and TR 38.857 [positioning]), link and system level simulations, etc.
- Additionally, use **field data** to *further* assess the performance *and robustness* in real-world environments. *How the field data set is obtained (per company vs. common data set) needs to be further discussed.*
- **KPIs** are broadly understood to be use case specific but a number of common metrics prevail, e.g., complexity and performance vs. proper (non-AI/ML) *state-of-the-art* baseline. *Similarly, overhead associated with enabling respective AI/ML scheme should be well documented and accounted for.*
- *Evaluated AI/ML schemes should be based on **offline training** for evaluation.*

Feedback Form 18: Companies inputs on Evaluation Methodology and KPIs for the candidate SI on AI/ML for Air-Interface

1 – Verizon UK Ltd

Agree with moderator's refinement based on intermediate round comments.

2 – Futurewei

Futurewei is ok with the first 3 bullets specified in the proposal. For the last bullet, we think it's too early to limit the evaluation based on "offline training" at this stage.

3 – Qualcomm Incorporated

We agree with the above refined proposals.

4 – China Mobile Com. Corporation

In our view, to make sure the performance comparison is fair, calibration is needed. Hence, we proposal to add a bullet:

Calibration based on a selected model is necessary.

5 – NTT DOCOMO INC.

We support the proposal.

6 – Apple Hungary Kft.

Apple: Agree with the proposal.

7 – LG Electronics Inc.

Re the last bullet, we cannot always assume that offline training shall be completed before applying parameters derived from AI in real network. We prefer to take training aspect into account as a part of SI evaluation methodology and KPI.

8 – Intel Technology India Pvt Ltd

Support the proposal

9 – Guangdong OPPO Mobile Telecom.

we are general fine. Then what is the plan to collect data in field? If it is not common among companies, how could be it used for evaluation?

10 – ZTE Corporation

We are fine with this high-level proposal on evaluation methodology and KPIs.

11 – Deutsche Telekom AG

We are generally fine with the proposals and the listed open issues.

With respect to use of field data, it is unclear from description if they should be used for training (incl. validation) or inference purposes or for both.

12 – SHARP Corporation

We support the proposal in general.

13 – Spreadtrum Communications

Support the proposal

14 – CAICT

We are generally fine with the proposals. The use of field data needs further discussion. It might not necessary to preclude that filed data is explicitly included during the SI phase for performance verification/calibrations between different companies.

15 – Fujitsu Limited

Regarding to **field data,** besides **'***per company vs. common data set***',** the methodology on how to use field data in the evaluation of AI models should be studied. For example, whether it is for training AI or only for testing AI or both?

16 – Fraunhofer IIS

The bullet "Additionally, use **field data** to *further* assess the performance *and robustness* in real-world environments. *How the field data set is obtained (per company vs. common data set) needs to be further discussed.*" is unclear to us. What kind of field data is foreseen and how would it be shared between companies? Some clarification is needed.

17 – NEC Corporation

We support the moderator's proposal.

18 – Lenovo (Beijing) Ltd

Agree with moderator's proposal. For bullet 4, it may be too strictly for the potential AI/ML schemes in this stage.

19 – Nokia Corporation

We agree to base evaluations on existing 3GPP scenarios. However, at least some of the selected scenarios shall include time-variant effects (e.g. time-variant traffic, user mobility) as it needs to be evaluated how well AI/ML schemes perform under such real-life conditions. Assumptions for training/exploration should be agreed as part of this SI. In our view we should not limit to the evaluations to offline training only. If companies provide methods on model recovery and online training to make sure that the performances are not impacted, it should be ok to consider in the evaluations as well. Overall, evaluation methodology discussions should be part of WG discussion and work. Stability, convergence, and scalability/generalization of AI/ML solutions are also very important aspects, so the new KPIs for monitoring this should be included in the SI scope. The actual KPIs should also be defined as part of the study item.

20 – Samsung Electronics Polska

We prefer a basic common dataset of field dataset for validation and how to collect it should be discussed first before assessing the performance and robustness in real world.

We don't support evaluated AI/ML schemes are only based on offline training. This should depends on the use case. And we strongly suggest to add at least one use case for online training as well.

On the other hand, for the uses cases with offline training, offline training for evaluation could be the first step for evaluating simulation-based scheme and algorithm performance.

21 – vivo Mobile Communication Co.

We prefer to update the part for field data as following. We are fine with other three bullets at least for evaluation purpose.

- Additionally, ~~use **field data** to *further* assess the performance *and robustness* in real-world environments.~~ ***whether and*** *how the field data set is obtained (per company vs. common data set) needs to be further discussed.*

22 – MediaTek Inc.

We agree with the above proposals.

23 – VODAFONE Group Plc

Agree. In addition to Universally agreed set o KPIs, we need to be able to specify or add operator-s[specific KPIs

24 – Ericsson LM

We don't support the moderator's proposal to agree on to use field data as there are several problems associated with it and we don't see the need for field data. Hence, the second bullet should be removed, or alternatively we suggest rephrasing the second bullet as follows: "*Additionally, whether to use field data to further assess the performance and robustness in real-world environments needs to be further discussed. How the field data set, if needed, is obtained (per company vs. common data set) needs to be further discussed.*"

Other bullets in the proposal are ok.

25 – HUAWEI TECHNOLOGIES Co. Ltd.

The last bullet on offline training should not be included at this stage. It seems different companies may have different interpretation/understanding on online training and offline training, therefore it is better to do more study/discussion in the study phase.

26 – Sony Corporation

We generally agree with the moderator summary.

We see both advantages and difficulties in using field data, but these issues can be discussed, as per the moderator's suggestion.

Some use cases may benefit from online training, including federated learning. The SI shouldn't restrict itself to offline training.

27 – InterDigital

- We think that the existing 3GPP channel models should be used as baseline. However further discussions are required on the need/feasibility/logistics of field data collection.
- The last bullet on '*Evaluated AI/ML schemes should be based on **offline training** for evaluation.' – is not clear to us.*

28 – ROBERT BOSCH GmbH

Agrees on the proposal

3.4 UE and Network involvement for AI/ML for Air-Interface SI

AI/ML for Air-Interface may mean different things to different companies. The moderator intent is to frame the discussion under a common language. Clearly, different use cases may require different collaboration frameworks. As a result, it would be good to have an agreed upon categorization for a more efficient and productive discussion going forward.

Based on some of the comments received, the Moderator has updated the various collaboration frameworks as follows (0->0a, 0b added, 3 removed):

0a) No collaboration framework: AI/ML algorithms purely implementation based and not requiring air-interface changes.

0b) No collaboration framework with modified Air-Interface catering to efficient implementation-based AI/ML algorithms purely implementation based.

1) Inter-node assistance to improve the respective nodes AI/ML algorithms. This would apply to UEs getting assistance from gNBs (for training, adaptation, etc.) and vice-versa. This level does not require model exchange between network nodes.

2) Joint ML operation between UEs and gNBs. This level requires AI/ML model instruction or exchange between network nodes.

While clearly Cat 0a does not require air-interface changes, it may provide an insightful baseline to assess the (performance/complexity) value of Cat 0b, 1, 2 based solutions and it is Moderator's understanding that should not be removed from the Categorization.

Question: Is there a need (from air-interface perspective) to distinguish whether the collaboration framework relates to training phase and/or inferencing phase?

Feedback Form 19: Companies' inputs on revised collaboration frameworks (incl. providing example use cases for each case) and answers to the Question above.

1 – Futurewei

Futurewei thinks it's too early to discuss the definition of collaboration levels at this stage as alignments are needed across companies. We suggest this discussion to be carried as part of the SI.

2 – CATT

Our view is that collabortive training between gNB and UE is too complex and premature to be included. Including collaborative training would make the scope too board and lead to nothing in the end.

3 – Qualcomm Incorporated

We agree with the categorization, but we prefer having the above categorization based on inference only. That is,

Cat 0b does not require any signaling for inference.

Cat 1 requires assistance information for inference.

Cat 2 involves over-the-air transmission of the inference output.

Regardless of which category the given AI/ML algorithm belongs, it's possible to define assistance data for training. Therefore, the collaboration framework for training and inference are two separate questions. In this respect, we propose to rename Cat 0a, 0b, 1, 2 into InfCat 0a, 0b, 1, 2 to make it clear that the categorization is based on collaboration framework for inference.

Similarly, we propose to define training-based categories as follows

TrainCat 0) No collaboration framework is needed for training

TrainCat 1) Inter-node assistance for training AI/ML algorithm. This could be in the form of assisting training data from gNB to UE, or vice-versa.

TrainCat 2) AI/ML model exchange between network nodes during training. Federated learning is one such example.

So, as an example, TrainCat0*InfCat1 refers to AI/ML algorithm framework needing assistance info for inference but no assistance data for training, and TrainCat1*InfCat0b for AI/ML algorithm framework needing no assistance info for inference but utilizing assistance training data for online training.

4 – China Mobile Com. Corporation

We think it is better to distinguish the collaboration framework in training phase and inferencing phase, respectively. In some cases, maybe the training is done at gNB/UE only, but for inferencing, the operation between UEs and gNBs is needed.

5 – NTT DOCOMO INC.

□We think it is better to keep 0a) in the list for performance reference used to compare with 0b), 1), and 2). Also, even if implementation-based AI/ML algorithms do not require air-interface changes to acquire inference of AI/ML, air-interface might need be modified to exploit the inference from AI/ML. For example, if implementation-based AI/ML can predict the channel or/and optimum beam, it is beneficial to be able to perform predicted beam indication (e.g., indicate beams with some time offsets to apply). Likewise, implementation-based AI/ML algorithms are worth studying for enhancements and even 0a) has potential of air-interface changes.

□Regarding the question, it is better to clarify collaboration frameworks in training phase or inference phase, because mixing them up could cause the misunderstanding between companies.

□The difference between 0b) and 1) are confusing to us. Especially, the definition of assistant data in 1) is not clear. Could you elaborate this point?

6 – Apple Hungary Kft.

Apple: Based on the last bullet of proposal in 3.3, "*Evaluated AI/ML schemes should be based on offline training for evaluation*", it seems all the collaboration framework 0a, 0b, 1 and 2 apply to inferencing phase only. It is not clear how this collaboration framework is related to offline training. For 1), suggest removing training if we agree proposal 3.3. "Inter-node assistance to improve the respective nodes AI/ML algorithms. This would apply to UEs getting assistance from gNBs (for ~~training~~, adaptation, etc.) and vice-versa. This level does not require model exchange between network nodes."

Further clarification of 0b) is needed. Hard to see the difference between (0b) and (1).

7 – LG Electronics Inc.

Updated frameworks look better organized than the previous version. Having Cat0a/0b is fine to us since AI can be an enabler for advanced NW/UE (e.g. NW/UE capable of channel prediction, etc.) while AI may not be visible in specification. But in general, we still think that this needs to be a part of SI including the question raised by moderator, i.e. no need to explicitly limit certain category in the SI objective.

8 – Intel Technology India Pvt Ltd

we think the categorization as described above is sufficient with the previous proposal that evaluation is based on offline training.

9 – Guangdong OPPO Mobile Telecom.

we are fine with 0b),1) and 2). The principle 1 is AI/ML algorithm will be a "black box". For 0a),if there is nothing additional exchange for that, how can It "provide an insightful baseline to assess the (performance/complexity) "?

10 – ZTE Corporation

We are generally fine with the four categories, but the difference between Cat 0b and Cat 1 may need further clarification. For example, one question is whether Cat 1 includes the case that gNB or UE provides input of the AI/ML model in the other side, while the AI/ML model itself keeps unchanged during inference. Or this inter-node exchange falls in Cat 0b? If such exchange falls in Cat 0b, we can further revise Cat 1 as

1) Inter-node assistance to improve the respective nodes AI/ML algorithms <u>with potential online change of AI/ML model parameters</u>. This would apply to UEs getting assistance from gNBs (for training, adaptation, etc.) and vice-versa. This level does not require model exchange between network nodes.

Then the different between Cat 0b and Cat 1 is Cat 0b does not require online change of AI model, but Cat 1 requires such online change.

Following this definition, example use cases for different categories can be

- Cat 0b: RS overhead reduction, positioning, beam management
- Cat 1: beam management
- Cat 2: CSI compression feedback

On the question from moderator, we think it is not needed to distinguish the training and inference procedures at least for now. As the training procedure will be based on offline training, training shouldn't impact air interface. In fact, we don't think to discuss too many details is helpful at current stage. We can

just agree on clear definition of these high-level categories, and further discussion on how to apply them in different use cases or procedures can be left in the SI for WGs.

11 – Deutsche Telekom AG

We think that there are more discussions needed to describe the cases in the collaboration framework.

As it was stated in 3.3 ”Evaluated AI/ML schemes should be based on offline training for evaluation” we would assume that only online training is considered here in addition to model inference. This is not clear from descriptions.

Cases (0a) and (0b) should be out of scope of the study as they can be based on present specification (0a) or modifications required for air interface can be considered in usual RAN(1) business.

(1) talks about ”Inter-node assistance”, but seems to be more on UE assistance for NW nodes. It is therefore not fully clear for us what is meant here.

Therefore, it is too early for us to agree already now to that details.

12 – SHARP Corporation

From specification impact point of view, we think there is no need to distinguish, because the collaboration level should be finally determined as the higher one between the collaboration level for training phase and the collaboration level for inferencing phase.

13 – Xiaomi Communications

Xiaomi : In our view, for inference phase, categorization may be needed. But we think it is too early to discuss this issue. This can be treated as one metric when we assess one scheme.

14 – Fujitsu Limited

This issue can be a part of ‘high level principles’ of AI for air interface, and to reach a common understanding among companies to the AI framework, functionalities, workflow in Rel-18.

Regarding to the question, we think there is a need to clarify collaboration frameworks in training phase and/or inference phase in order to avoid misunderstanding to the scope.

15 – CAICT

The collaboration level is a general categorization from the point of view of information exchanging between gNB and UE. Based on the discussions, there is no clear definition of training phase and inferencing phase and their relationship. The training mechanisms for each case should be part of SI.

16 – NEC Corporation

We share the same view with the moderator that cat 0a can work as the baseline for the result comparison. For the question, we don’t see the need because the classification of the collaboration frameworks is in terms of specification changes, instead of AI/ML models. The models including training and inference can be put into different categories considering different collaboration requirements.

17 – Lenovo (Beijing) Ltd

We think the UE and Network involvement issue is not only the use case specific, but also related with the AI/ML approaches per use case, and different phases, e.g., training and phase. Thus, we suggest focusing on Cat.0a/b in the SI for air interface, and leaving Cat.1/2 to higher layer, e.g., SI for NG-RAN.

18 – Nokia Corporation

In overall the moderator's list is ok. However, depending on the use case, company proposals on how to apply AI/ML may differ. When comparing these different proposals, the collaboration framework should be properly accounted in the comparisons. Therefore, in some form, it would be better to integrate assumed framework in the evaluation methodologies. In our view details should be discussed as part of the study in WGs.

19 – Samsung Electronics Polska

First of all, we think the above questions should be studied in SI phase.

For the use cases with online training, we should not distinguish training phase or inferencing phase.

For the use cases with offline training, we agree with the categorization for inference phase. For training phase, data may be collected and model be trained offline in dependence of inference phase. We could first focus on and standardize for offline training solution at current stage. It is too early to conclude the categorization for both training and inferring phase.

20 – MediaTek Inc.

We have similar view as Nokia. We are fine with the list, but not sure why it has to be decided now, and consider it should be part of the SI, i.e. for each use case, options of collaboration framework should be studied and evaluated.

21 – VODAFONE Group Plc

At this stage we could only agree to 0a , also we need to consider the complexity of extracting low layer information from the network hardware as this could put sever load on the equipment, as additional circuitry would be required to 'Fetch' and 'Store' data

22 – Ericsson LM

Ericsson's understanding of this collaborative framework is that it follows the higher-level principle that it is built around implementation-based AI/ML algorithms, such that any potential specification activities will be on interfaces and signalling.

These collaboration frameworks are a good starting point, but we believe we need to further discuss the definition of each of them, since companies likely have different interpretation at this stage. To progress, we share our current understanding of these four frameworks below:

Our understanding of Framework 0a is that it only concerns "inference mode" of fixed AI/ML models (e.g., neural networks with fixed weights and biases) within the existing standard (no air interface changes). We also understand that the AI/ML models are proprietary / implementation-based algorithms

Our understanding of Framework 0b is that it only concerns "inference mode" of fixed AI/ML models (e.g., neural networks with fixed weights and biases), together with the required new signaling, new reference signals etc, to trigger use of the "inference mode" in the UE and/or gNB. We also understand that the AI/ML models are proprietary / implementation-based algorithms

Our understanding of Framework.1 is that it includes "node category exchange", where information about the UE's chipset / AI/ML capabilities is exchanged together with other assumptions related to channel,

RS, QCL etc. To avoid any potential confusion, we suggest that "inter-node assistance" is replaced by "gNB-UE exchange of assistance information". We also understand that the AI/ML models are proprietary / implementation-based algorithms.

Our understanding of Framework.2 is that it involves joint optimization of AI/ML algorithms at both the gNB and UE side (when applicable, e.g. CSI enhancements). These algorithms are proprietary, either fully or partially.

On the Question: This will likely depend on use case and the particular AI/ML solution. However, there could be a need (from an air-interface perspective) to distinguish whether the collaborative framework relates to *training*, *data collection*, and *inference* phases

23 – HUAWEI TECHNOLOGIES Co. Ltd.

Our view is that it is too early to identify the inter-node collaboration levels at this stage. As we commented in the previous rounds, UE and Network involvement for AI/ML for Air-Interface should be discussed for each specific use case, by considering the procedures of model training and model inference. The signaling between network and UE can also be different for different use cases. The corresponding details need to be ironed out during the study item. We understand the intention from the moderator, however without more detailed discussion on the use cases, it is difficult to judge whether the current categorization is appropriate or not, which may result in wrong direction also. Therefore, we still think this kind of discussion can be done in the study phase.

24 – vivo Mobile Communication Co.

We have the following comments for the categories on collaboration levels:

- Our current understanding for different categories is that Cat 0a) does not require additional specification, Cat 0b) requires some air interface enhancement, but the exchanged information is not directly related to inference, Cat 1) requires additional information exchange, but the exchanged information is directly related to inference, Cat 2) requires model exchange between two sides.
- Further split between training and inference for above categories is helpful to align companies understanding. However, we also sympathize with the comment that this may belong to the study item issue, e.g., it can be discussed during pre-Study phase.
- Cat 0a) is necessary to be listed as the reference to define baseline performance.
- Having these four categories as listed example category is helpful for study phase discussion.

25 – Sony Corporation

We don't understand the sentence for framework 0b. Can this please be re-phrased?

We are generally OK with the categorization.

The collaboration framework can be applied to both the training and inference phase, or to joint training-inference, depending on the AI/ML scheme applied.

26 – InterDigital

- We think that categorization framework should be generic enough to support all (/future) use cases. So we support the need to distinguish the categorization framework between training and inference.

27 – ROBERT BOSCH GmbH

Joint ML operation between UEs and gNBs may be very difficult to start with in Rel-18. However, we support the Inter-node assistance (between UE and gNB) to improve the respective nodes AI/ML algorithms.

28 – KDDI Corporation

We support views from Futurewei and also think that before discussing the this topic whether to add new use cases needs to be discussed.

4 Conclusions

The Moderator thanks all the contributing companies for the comments received during the three discussion phases.

While all Moderator's proposal in the Final round received a very broad support. There are still some company comments not fully aligned with what has been proposed.

As a result, it is the Moderator's understanding that there is no need to endorse any proposals at this point. However, from the discussions, we have gained a much clearer understanding on the following areas:

- Possible project structure for AI/ML related projects in Rel-18, including AI/ML for NG-RAN and AI/ML for Air-Interface

For a candidate Rel-18 SI on AI/ML for Air-Interface, the following areas have been discussed at length and we have a much better understanding now which will be helpful for the drafting of a potential SID:

- Use cases of interest for candidate Rel-18 SI on AI/ML for Air-Interface
- Evaluation methodology and KPIs
- UE and Network involvement including various degrees of collaboration between participating nodes

In that sense, we have a good understanding about from where to pick up things in the next round of discussions.