3GPP TSG-RAN WG2 #121 R2-23XXXXX

Athens, Greece, February 27 – March 3, 2023

Agenda Item: 8.16.1

Source: Ericsson, vivo

Title: Outcome of [Post120][054][AIML18] Data Collection (Ericsson / vivo)

Document for: Discussion, Decision

# 1 Introduction

The intention of this document is to invite companies to share their views regarding *data collection* aspects for the SI on “AI/ML for NR Air Interface”. Taking these into account, the Rapporteur of the discussion provides a set of proposals to be further discussed during RAN2#121.

The scope of the discussion is given by the following email thread:

|  |
| --- |
| * [Post120][054][AIML18] Data Collection (Ericsson / vivo)   Scope: Long email discussion for next meeting, on data collection (focus on monitoring and training), on to what extent existing methods can be useful including also identifying these existing methods and their potential extensions  Intended outcome: Report  Deadline: Long |

**Deadline for comments: Friday Feb 10th, 2023, 1000 UTC**

**Inactive periods:**Dec 23 – Jan 6 Jan 23 – 27

Below you can find the list of participating companies and their respective responsible delegates.

|  |  |  |
| --- | --- | --- |
| **Company** | **Delegate name** | **Email address** |
| Apple | Peng Cheng | pcheng24@apple.com |
| OPPO | Jiangsheng Fan | fanjiangsheng@oppo.com |
| Qualcomm | Rajeev Kumar | rkum@qtu.qualcomm.com |
| Lenovo | Congchi Zhang | zhangcc16@lenovo.com |
| Interdigital | Oumer Teyeb | oumer.teyeb@interdigital.com |
| vivo | Boubacar Kimba D.A. | kimba@vivo.com |
| Xiaomi | Xing Yang | Yangxing1@xiaomi.com |
| NEC | Hisashi Futaki | hisashi.futaki @ nec.com |
| LGE | Soo Kim | soo.kim@lge.com |

# 2 Discussion

As per the Email Discussion description above, the intention of this document is to touch upon data collection for:

* model monitoring and,
* model training.

For this, RAN2 should then analyse whether existing methods could be (re)used or extended. Or whether there is a need for new methods.

On the above, the Rapporteur suggests considering the use cases under study, to later focus on their requirements and architecture-related aspects.

## 2.1 Use cases

The three different RAN1-agreed use cases and their respective sub use cases are listed below:

1. CSI feedback enhancement
   * Spatial-frequency domain CSI compression using **two-sided AI model**
   * Time domain CSI prediction using **UE-sided model**
2. Beam Management (BM) enhancement
   * Spatial-domain DL beam prediction, with **one-sided AI model** (i.e., either in UE or NW)
   * Temporal DL beam prediction, with **one-sided AI model** (i.e., either in UE or NW)
3. Positioning accuracy enhancement
   * Direct AI/ML positioning
     + *Note: this refers to the fact that the AI/ML model is directly producing the UE location as output*
   * Assisted AI/ML positioning
     + *Note: this refers to the fact that the AI/ML model is producing an existing or new measurement report that is used to estimate the UE location using legacy positioning methods (e.g., triangulation).*
   * For the above 2 points (i.e., direct/assisted AI/ML positioning), RAN1 have captured the following (sub)cases:
     + Case 1: UE-based positioning with **UE-sided model**, direct AI/ML or AI/ML assisted positioning
     + Case 2a: UE-assisted/LMF-based positioning with **UE-sided model**, AI/ML assisted positioning
     + Case 2b: UE-assisted/LMF-based positioning with **LMF-sided model**, direct AI/ML positioning
     + Case 3a: NG-RAN node assisted positioning with **gNB-sided model**, AI/ML assisted positioning
     + Case 3b: NG-RAN node assisted positioning with **LMF-sided model**, direct AI/ML positioning

Above, the Rapporteur have highlighted where the AIML model inference is located for each agreed use case, i.e., UE- or gNB/NW-sided AIML models.

In this sense, it seems important for to clarify the scope of RAN2’s discussion concerning the functionality-to-entity mapping within the network. This, since during RAN1#109-e the following was captured in the Session Notes (see [R1-2205695](http://www.3gpp.org/ftp//tsg_ran/WG1_RL1/TSGR1_109-e/Docs//R1-2205695.zip)):

|  |
| --- |
| Observation  Where AI/ML functionality resides depends on specific use cases and sub-use cases.  Conclusion   * RAN1 discussion should focus on network-UE interaction.   + AI/ML functionality mapping within the network (such as gNB, LMF, or OAM) is up to RAN2/3 discussion. |

On the above, the Rapporteur would like to start by stressing that RAN2 should limit their analysis and discussion to RAN2-specific aspects.

Clearly, the positioning (sub) use cases are likely to lead into discussions that may involve the LMF and, perhaps, other entities outside of RAN. This aspect is addressed further below in the document (see Section 2.2.2). However, for the other cases, it seems reasonable to start by assuming that RAN2 should focus on data collection information that is terminated in the gNB, LMF or UE.

Furthermore, the Rapporteur would like to continue by stressing that RAN3 does not have Time Units (TUs) allocated to this SI. Therefore, it is even more important to limit the scope of the functionality-to-entity mapping discussion (and here not only for data collection aspects) to RAN2 domains of expertise and not to that of other WGs.

1. RAN2 should not rely on RAN3 regarding architecture or functionality-to-entity mapping matters, since RAN3 does not have TUs assigned for this SI.

Further developing the above thread, the Rapporteur observes that UE-sided models based on data collection at the UE side have minor or no implications to RAN2 protocols, as a lot is proposed by companies to be left to implementation. As for these cases, the UE seems already capable of collecting data (e.g., by UE implementation). Indeed, the Rapporteur understands that some signalling, or specific configuration could eventually be needed by UEs for data collection purposes in UE-sided AIML models. However, for these cases it is our understanding that the UE does not need to report the collected data to the gNB. Hence, UE-based data collection for UE-sided AIML models appear to mostly be a RAN1 issue right now.

1. For UE-sided AIML models, UEs do not necessarily need to report the collected data for model training to the gNB/LMF.
2. RAN1 can later provide requirements (e.g., related to configuration, signalling, etc.) which could facilitate the process of UE data collection for UE-sided AIML models.

On the contrary, gNB/LMF-sided AIML models need to collect data gathered by UEs to monitor and train the models. This data may be UE measurements and performance metrics that are configured and reported to the gNB/LMF. One point which needs to be discussed, for example, is to which extent the characteristics of the data to be collected differs from existing measurements the UE may already be able to perform and report to the gNB. Hence, the Rapporteur observes that there is perhaps a need for RAN2 to start by focusing on gNB/LMF-sided AI/ML models.

1. There seems to be a need for RAN2 to start by focusing on gNB/LMF-sided AIML models. Since for these cases, the UE needs to be configured to report collected data (e.g., measurements, performance metrics) to the gNB/LMF.

The Rapporteur would like to understand companies’ views regarding the above. Hence the following question.

**Q1)** Based on the above: Do you agree that RAN2 should start discussing the solutions and specification impact analysis centered around UE measurements/reporting and data collection for gNB/LMF-sided models?

|  |  |  |
| --- | --- | --- |
| **Company** | **Answer** | **Comments** |
| Apple | No | First, we disagree observations 2-4 can lead to the proposal in Q1 (i.e. RAN2 can start the study UE reporting for NW-sided model):   1. Observation 2/3:We think both "UE-sided model" and "gNB/LMF-sided model" need to define measurement requirement because both models need UE's measurements. The difference is whether to define requirement for UE reporting (i.e. "UE-sided model" may not need to define reporting requirement). For UE measurement, both RAN1 and RAN2 have spec impacts to be analysed. 2. Observation 4:We think whether impacts on UE measurement/reporting should be studied by RAN1 or RAN2 depends on the related measurement is L1 measurement or L3 measurement. Even for "gNB/LMF sided model", if the related measurement is L1 measurement, RAN1 should first study it. Because CSI measurement, BM and measurement for positioning are all L1 measurement, we are not sure why the study should start from RAN2?   Secondly, we think Rapporteur seems to miss some aspects of data collection study. It is our understanding the on-going study on "data collection" in RAN1 include the following 4 parts:   1. The assistance signalling from gNB to UE on data collection 2. The assistance signaling from UE to gNB on data collection 3. The delivery of dataset from UE to gNB 4. The delivery of dataset from gNB to UE (being discussed in AI/ML based positioning)   Rapporteur only focus on c) but missed 1), 2) and 4). We think it is not correct. **We want to confirm that all above 4 aspects are in RAN2 scope of data collection. We see no reason to preclude/deprioritize any one of them at this stage without any RAN2 study.**  Thirdly, we have concern that the requirements for data collection is not clear at this stage. We are not sure how RAN2 can make progress without clear requirement? Thus, **we suggest RAN2 can currently confirm the scope but need to wait sufficient RAN1 input on data collection requirement before detailed data collection study.**  Finally, our suggestion on RAN2 study of data collection can be summarized as:  1) Confirm and agree RAN2 scope of data collection study.  2) Agree what kinds of requirements need to be considered in RAN2 data collection study and the requirements may be use case specific.  3) Then, wait sufficient RAN1 inputs on details of requirements.  4) Based on agreed requirements for AI/ML data collection, RAN2 can discuss whether to define a new data collection framework or extend existing framework.  5) As the final step, RAN2 specify detailed signalling and procedure of UE data collection. |
| OPPO | No | We think it’s too early to make such assumption. If we check RAN1 discussion, we will find that only high level agreements were made for data collection. For general part, RAN1 made the following agreement:  *Conclusion*  *Data collection may be performed for different purposes in LCM, e.g., model training, model inference, model monitoring, model selection, model update, etc. each may be done with different requirements and potential specification impact.*  *FFS: Model selection refers to the selection of an AI/ML model among models for the same functionality. (Exact terminology to be discussed/defined)*  For use case specific topics, data collection requirements are quite different, no much progress was made so far.  Based on above, it’s hard to say that RAN2 should focus on network side AI model, maybe this scenario is typical for model training, but UE side AI model training or monitoring may also involve network configuration, e.g. RS configuration or performance metrics, so it’s safer for RAN2 to leave the discussion open for now.  Another thing is that it’s better to split the discussion between model training or model monitoring as the wanted data types as well as data size requirement are quite different. Maybe a general framework is possible in the future, but we can’t assume a unified data collection framework is needed at the beginning without clear evaluation. |
| Qualcomm | No | We do not agree with observations 2-4. We have the following understanding,  Observation 2/3: Similar to Apple's comment, we believe both the UE-sided model and gNB/LMF-sided AIML models will require data collection for training, inference, monitoring, update, etc. Furthermore, our understanding is that data collection should be discussed for different purposes and limitations of existing data collection frameworks in meeting them.  Observation 4: For inference and monitoring, RAN1 is discussing different required parameters. For training, RAN2 may start discussing the limitations of existing methods. We believe that for training and model development, large-scale data collection may be required.  Furthermore, we agree with Apple's comment on the requirements of assistance signaling from gNB to UE for data collection.  In our understating, we should   1. Understand the requirements of data collection for training, inference, monitoring, update, etc. 2. Evaluate the limitations/benefits of current data collection methods for training, inference, monitoring, updating, etc. 3. Study the requirements for new data collection methods   We believe that we should start evaluating the existing methods and requirements for a new framework. Our understanding is that until the last meeting of RAN1, RAN1 will keep evaluating the parameters required. We believe that RAN2 should wait for the RAN1 study on the required parameters for different use cases and focus on the data collection framework for supporting model training, inference, monitoring, updating, etc. |
| Lenovo | No | We also don’t agree with Observation 2 to 4.  *“Observation 2 For UE-sided AIML models, UEs do not necessarily need to report the collected data for model training to the gNB/LMF.”* implies that the UE-sided AIML model will not be trained by gNB/LMF, wherein UE measurements report can be used for training. In contrary, RAN1 discussion and many proposals in RAN2 (as well as the other RAN2 email discussion on model transfer delivery) still consider the scenario that a UE sided AIML model is trained and delivered from the NW.  We believe data collection for both UE sided model and gNB/LMF sided model shall be studied with equal priority.  We also agree with above companies that RAN2 can first study using existing data collection framework for the stages of AI operation i.e., AI training/inference/monitoring/update, even though this email discussion was agreed to focus on training and monitoring. |
| Interdigital | No | As the companies above have indicated, UE-sided models and 2-sided models could also require data collection (at least for training purposes, if the training is to be done on the network side). Also, we think that the main idea behind this email discussion is identifying the capabilities/limitations of current data collection frameworks and when RAN1 has made considerable progress regarding the requirements of data collection for the different use cases, RAN2 can be ready to reuse the current data collection framework (if they are capable) or discuss on how to update the existing framework or even design new ones to accommodate the requirements. |
| vivo | See comments | As RAN2 would not discuss observations online, we prefer not to focus on them. The controversial observations can be refined to include the UE-sided model.  However, considering data collection from UE to gNB/LMF, as a first step, may allow the discussion to be more focused on some potential RAN2 agreements. But, this does not exclude other data collection, e.g., data collection from UE to other Network entities or data collection from network to UE. So, we think the discussion on UE measurements/reporting and data collection for gNB/LMF-sided models can be considered as an INITIAL baseline, which can facilitate the discussions for other scenarios, e.g., UE-sided AIML model is trained and delivered from the NW.  Additionally, RAN2 may also consider other data collection aspects from UE to other network entities or from other network entities to UE. If the discussion on those data collection aspects involves other WGs, such as RAN3 or SA2 (which have no TU for this SI), RAN2 may decide to leave out those detail discussions until potential WI phase and involve them, if necessary. |
| Xiaomi | Comments | We understand the data can be used for model training, inference and performance monitoring.  We understand the so called one side AIML means the inference is done at one side. So, even in one-sided AIML models, it’s possible the training and performance monitoring are located in the other side, which may also require data collection and data exchange. We suggest to focus on the functionality required for data collection. So, we suggest following modification,  Do you agree that RAN2 should start discussing the solutions and specification impact analysis centered around UE measurements/reporting for data collection for AIML model training and performance monitoring at gNB/LMF |
| NEC | No | As many companies commented above, the observations 2-4 seem restrictive or a bit biased. We also assume that the data collection for UE-sided model needs to be studied as well as gNB/LMF-sided model, and they may have different requirements.  In general, we tend to agree with the summary 1-3 from Qualcomm. |
| LGE | Comments | As other companies have mentioned above, we disagree with O2-O4 as the network can train/monitor the model for the UE side model.  We think it should be discussed separately depending on the purpose of data collection (e.g., model training and model monitoring) and use case. We also don't believe that network entities for training and monitoring are always the same.  From this perspective, So, we suggest the following modification,  Do you agree that RAN2 should start discussing the solutions and specification impact analysis centered around UE measurements/reporting for data collection for AIML model training and/or performance monitoring at gNB/LMF |

Following the previous reasoning, the Rapporteur would then like to see whether it would be possible to assume the following.

**Q2)** Would it be possible to agree on the following: RAN2 assumes that for gNB/LMF-sided AIML models, the data collection information is terminated in the gNB/LMF?

|  |  |  |
| --- | --- | --- |
| **Company** | **Answer** | **Comments** |
| Apple | No | This proposal may mistakenly preclude the SON/MDT framework because it is terminated in TCE rather than gNB/LMF, although we think it is not the intention of Rapporteur. Meanwhile, this proposal also precluded the option that termination in OAM, which is conflicted with RAN3 agreement.  Without any study in RAN2, we don't think RAN2 can make this assumption now. |
| OPPO | No | The suggestion from Rapporteur is only one possible way, as mentioned by Apple also, many existing data collection frameworks are precluded if we agree this proposal, so let’s focus on existing data collection framework first, whether the data collection information is terminated in the gNB/LMF can be considered later. |
| Qualcomm | No | As mentioned above, we believe that RAN2 should focus on the data collection framework for supporting model training, inference, monitoring, updating, etc. The proposal above may work for inference but not for training and monitoring.  We believe that with the above proposal in Q2, the model training and development cannot be supported. The collected data may not be i.i.d and any trained model for such data collection methods will be highly overfitted, i.e., models cannot be generalized.  We want to highlight the requirements for data collection for training, inference, monitoring, updates, etc. are different. Therefore we believe that RAN2 should study the requirements for data collection for training, inference, monitoring, updates, etc. first. |
| Lenovo | No | The proposal made by the rapporteur assumes the gNB/LMF-sided AIML models are always trained by gNB/LMF themselves. However, “where AI training happens” is still an open question which is also relevant to the discussion in the other RAN2 email discussion on model transfer/delivery. In addition, as companies above mentioned, at least the SON/MDT framework allows the data terminated at TCE/OAM. |
| Interdigital | No | We think this is not a problem/question that needs to be discussed at this time. What we should focus first is about current data collection frameworks and their capabilities/limitations. |
| vivo | See comments | For gNB/LMF-sided AIML models*,* the data collection information may be terminated in the gNB/LMF, in case the data collection and model training is done at gNB/LMF.  Besides, for the SON/MDT framework, the data collection information is terminated in the gNB from the RAN2 perspective as how the gNB further processes and forwards the data to the TCE is in SA5 scope.  So, we propose to consider as a potential revision of the assumption as follows:*“In case the model training is done at gNB/MLF*, RAN2 assumes that for gNB/LMF-sided AIML models, the data collection information can be~~is~~ terminated in the gNB/LMF*”*. |
| Xiaomi | Comments | We think it’s one of the solutions to terminate the data collection at gNB/LMF. But it may be premature to exclude other possibilities, e.g. data is further transferred to other entities. |
| NEC | See comments | We think this assumption is too early to be made, although this is one of possible way. Probably without this assumption, RAN2 can discuss and progress. |
| LGE | No | We also think it’s one of the solutions to terminate the data collection at gNB/LMF. Therefore, we cannot assume it for all gNB/LMF-sided AIML models. |

*[Rapporteur to add summary of views]*

1. To be added according to companies’ views…

## 2.2 Requirements

For RAN2 to design data collection solutions for the (sub)use case, one should first focus on the requirements needed for each. Given the questions above and since we now focus on model training and model monitoring, the Rapporteur considers that there are 4 main aspects to consider:

1. The content of the data
2. The data size (e.g., for model training)
3. Latency, periodicity, or “efficiency” requirements that could differentiate model training from model monitoring
4. Configuration-related requirements

Considering the different (sub)use case of this SI, one could possibly further extend the 4 aspects above as follows:

1. **The content of the data**
   * radio measurements
   * non-radio measurements, e.g., assistance data, monitoring metrics (e.g., accuracy, predicted outcome)
2. **The data size**
   * how many time instances are needed per UE or/and per cell?
   * the amount of UEs or/and cells needed in the data collection step to accurately train the model,
   * payload size of UE report.
3. **Latency, periodicity, or “efficiency”**
   * is the model performance monitoring a time critical matter?
   * are there latency requirement differences between types of training?
   * how often does the data need to be collected?
   * what is the validity period for data? (i.e., how quickly does it become outdated)
4. **Configuration-related requirements**
   * how to eventually configure a UE to e.g., measure, store, and report data
   * whether there is a need to have periodic or event-triggered data collection,
   * RRC-state linked to data collection/reporting
   * scenarios/conditions for which the data can be discarded

**Q3)** Do companies agree to start by focusing on data collection requirements related to: a) content of the data, b) data size, c) latency, periodicity, or “efficiency”, d) configuration-related requirements?

Note: Companies are also encouraged to provide comments on other aspects that need special attention at this stage of the SI.

|  |  |  |
| --- | --- | --- |
| **Company** | **Answer** | **Comments** |
| Apple | See comments | First, we think the discussion should differentiate between model training and data monitoring (i.e. inference). Rapporteur seem to mix them together or think they share the same requirement. But our understanding is their requirements may be different. For example, the data size for training should be much larger than data monitoring. And the latency requirement for training is much loose than data monitoring (i.e. inference).  Secondly, we think RAN2 can first agree that a) b) c) can generally be considered, but their details for each use case should be further studied. For d), our understanding is that it is a stage 3 signaling detail rather than requirement. So, we don't think d) should be included as one requirement.   * Here, we have a clarification question on c): what is "efficiency" means? our understanding is to reduce redundancy of data collection, but rapporteur please confirm.   Thirdly, before RAN2 study, we think RAN1 inputs on a) b) c) are required and maybe RAN4 input on c) is also required.  Finally, we think a missed requirement is security and UE privacy, which is an very important aspect.  As summary, we suggest to agree below 4 general requirements of data collection design, but their details should be further study (i.e. current sub-bullets need further study):   1. **The content of the data** 2. **The data size** 3. **Latency, periodicity, [or “efficiency”]** 4. **Security of data and UE privacy** |
| OPPO | See comments | As mentioned in Q1, it’s strange to mix the discussion between model training and model monitoring as the data collection requirements may be quite different, so better to split the discussion.  Then, in our understanding, RAN2 data collection study should focus on signalling and procedure based on clear data collection requirements, it’s obvious that bullet a)/b)/c) needs RAN1 inputs as only RAN1-led use cases are considered so far, if high layer-led use cases are introduced in the future, RAN2 can discuss a)/b)/c) without waiting RAN1 progress, but for now, we are not sure whether we can discuss the data collection requirements even before RAN1 for RAN1-led use cases.  Regarding bullet d), this is also pending on RAN1 requirements, just like CSI configuration, the signalling is designed by RAN2, but the requirements are coming from RAN1.  Based on above, we think it’s not suitable for RAN2 to discuss the data collection requirements while no clear progress was made in RAN1. |
| Qualcomm | See comment | In our view, data collection requirements for training, inference, and monitoring will be different. For training generally, large data collection and independent and identically distributed (non-correlated data) are desired.  We further believe that RAN1 is evaluating the aspects such as,   * The content of the data * The data size * Latency, periodicity, or “efficiency” * Configuration-related requirements   Therefore, we believe that RAN2 does not need to evaluate these aspects and wait for RAN1 progress.  RAN2 should evaluate if the   * Current data collection frameworks/methods are sufficient for training? * Current data collection frameworks/methods are sufficient for inference? * Current data collection frameworks/methods are sufficient for monitoring both real and non-real-time? * Current data collection frameworks/methods are sufficient for model updates? * Study requirements for new data collection methods based on requirements for training, inference, monitoring both real and non-real-time, updates, etc. |
| Lenovo | See comment | We also think the data collection requirements for training, inference, and monitoring, update could be different, thus shall be analysed case by case. For instance, the data collection for training could tolerate more “latency” than model update or monitoring.  So, generally we agree with studying these points but separately for each of the four categories that we mentioned above: mode training, model inference, model monitoring, and model update.  For a) b) c), and part of d), those are under the scope of RAN1 discussion, RAN2 can wait for more RAN1 progress before discussing the related RAN2 aspects. |
| Interdigital | See comment | We also think that RAN1 input is needed to identify the requirement for the different use cases as well as the different purposes of data collection (i.e., training, inference, monitoring, etc.). We could end up doing double/redundant work if we start trying to come up with the requirements in RAN2 while RAN1 is doing the same. Thus, as we have indicated above, we propose to limit the RAN2 discussion at this time to the identification of the capabilities/limitations of data collection frameworks. |
| vivo | See comment | We agree the requirements can be categorized in two classes: 1) data collection requirements for model training and 2) data collection requirements for model monitoring.  So, based on this assumption RAN2 can discuss which of the above requirement apply to 1) and/or 2) from RAN2 perspective. Any requirement which necessitates RAN1 involvement can be later considered based on RAN1 progress. |
| Xiaomi | Comments | In general, we are fine with the main bullets. But sub-bullets may need further discussion/clarification.  For example, we wonder whether validity of the data is needed. If data is used for training, the data can always be useful. And, it’s not clear to us why RRC-state is linked to data collection.  Also, it may be good to separate the discussion for the data collection of monitoring and training, since the requirement for monitoring and training may be different. |
| NEC | See comments | We assume that RAN2 can start to confirm these 4 aspects (a-d) need to be studied. When studying, the data collection for model training and model monitoring should be separately discussed, as the requirements could be different as companies above already commented.  Then, before going into details e.g. use case specific aspects, RAN2 should wait for further RAN1 inputs. |
| LGE | Comments | Since a)-d) is also discussed in RAN1, we think RAN1 input is needed to prevent duplicated discussion.  Also, the content may vary depending on whether the data collection is training, inference, or monitoring.  RAN1 also discusses separately. For example, in the case of beam prediction, L1 reporting enhancement is considered for AI/ML model inference. For Trainig, spec impact (how to initiate data collection) and configuration are being discussed. For monitoring, for beam measurement based on a set of beam, RRC-based/L1-based content is being considered.  Therefore, RAN2 should study what and how to add to existing procedures or extensions of existing procedures according to input from RAN1. For this, it is necessary to first analyze the existing methods and procedures depending on the purpose of data collection. |

*[Rapporteur to add summary of views]*

1. To be added according to companies’ views…

### 2.2.1 Beam Management enhancements

The Rapporteur acknowledges that RAN1 have not provided explicit input to RAN2 concerning requirements for the Beam Management use cases. However, by focusing on NW-sided AIML models and considering the above “initial requirements”, the Rapporteur understands that the following could be considered to design solutions:

1. **The content of the data**
   * Radio measurements:
     + L1-RSRP measurements and/or SSBRI/CRI (beam ID) of a set of beams at one- or multiple-time instances
   * Non-radio measurements:
     + Cell ID, area ID, carrier frequency, UE/NW antenna beam configuration/ID, time stamp, UE location/mobility/rotation, measurement accuracy, or measurement resolution, etc…
   * Monitoring metrics:
     + Prediction accuracy
2. **The data size**
   * Data should be collected from multiple UEs, at one- or multiple-time instances
   * The number of bits needed for reporting measurements per time instance depends on the number of beams a UE is configured to measured and report
3. **Latency, periodicity, or “efficiency”**
   * As per RAN1#111’s agreement:
     + For NW-sided AIML model monitoring, study the necessity and potential spec. impact from the following aspects:
       - UE reporting of beam measurement(s) based on a set of beams indicated by gNB,
       - Signaling, e.g., RRC-based, L1-based
4. **Configuration-related requirements**
   * It should be possible to configure a UE to store measurements on multiple occasions and then report the accumulated data to the NW
   * Periodic and event-triggered data collection approaches could be considered for further study

**Q4)** Companies are invited to comment on the above requirements, and whether these could be considered for designing data collection solutions for NW-sided beam management use cases.

Note: Companies are encouraged to provide further views or other aspects that should be considered.

|  |  |
| --- | --- |
| **Company** | **Answer / Comments** |
| Apple | See our comments to Q3. For now, we don't think RAN1 details on a) b) c) related to BM are clear. Thus, RAN2 should wait further RAN1 input. |
| OPPO | As mentioned in Q1, no much progress was made in RAN1 for data collection so far for use case specific topics, so our intention here is to go many steps further than RAN1? If that is the case, we don’t think it’s feasible for RAN2 to do this. |
| Qualcomm | See comment for Q3. |
| Lenovo | Same comment as Q3, and we understand the sub-bullets under a)b)c)d) are given only as examples, since RAN1 is still discussing, e.g., the exact input data for training, and data size. |
| Interdigital | Same comment as Q3 |
| vivo | As commented in Q3, RAN2 can discuss which of the above requirement apply to 1) data collection requirements for model training and 2) data collection requirements for model monitoring from RAN2 perspective. Any requirement which necessitates RAN1 involvement can be later considered based on RAN1 progress. |
| Xiaomi | Same as Q3. |
| NEC | See comment for Q3 |
| LGE | Same as Q3 |

*[Rapporteur to add summary of views]*

1. To be added according to companies’ views…

### 2.2.2 Positioning accuracy enhancement

The Rapporteur acknowledges that RAN1 have not provided explicit input to RAN2 concerning requirements for the positioning use cases.

Now, as per what has been described in Section 2.1, if we limit the scope of the current discussion to gNB/LMF-sided AIML models, then, only Case 2b (i.e., UE-assisted/LMF-based positioning with LMF-sided model, direct AI/ML positioning) and 3a (i.e., NG-RAN node assisted positioning with gNB-sided model, AI/ML assisted positioning) should be considered for this email discussion.

Starting with, Case 3a, we observe that arguably this seems not to need/include UE reporting data to the gNB. As for this case, the UE the gNB can perform measurements for model training/monitoring according to Sounding Reference Signals (SRS) in the uplink.Taking this and the above “initial requirements” in consideration, the Rapporteur understands that the following aspects could potentially be considered to design solutions:

1. **The content of the data**
   * Radio measurements:
     + Channel Impulse Response (CIR) information
   * Non-radio measurements:
     + UE location (ground truth label, which can be obtained from LMF)
2. **The data size** 
   * Data should be collected from multiple TRPs, at one- or multiple-time instances
   * The number of bits needed for collecting a single channel measurement (CIR) depends on how the channel is represented, i.e., the data format and preprocessing/quantization
3. **Latency, periodicity, or “efficiency”**
   * No latency requirements for training, unclear for monitoring
4. **Configuration-related requirements**
   * Existing configuration seems to fulfil the requirements (i.e., UE is configured to transmit SRS for gNB to perform SRS-based channel measurements)

**Q5)** Companies are invited to comment on the above requirements and whether these could be considered for designing data collection solutions for NG-RAN node assisted positioning with gNB-sided model?

Note: Companies are encouraged to provide further views or other aspects that should be considered.

|  |  |
| --- | --- |
| **Company** | **Answer / Comments** |
| Apple | See our comments to Q3. For now, we don't think RAN1 details on a) b) c) related to AI/ML based positioning are clear. Thus, RAN2 should wait further RAN1 input. |
| OPPO | The similar view as Q4. |
| Qualcomm | See comment for Q3. |
| Lenovo | Same comment as Q3, and we understand the sub-bullets under a)b)c)d) are given only as examples, since RAN1 is still discussing, e.g., the exact input data for training, and data size.  To be specific, the radio measurements only include CIR in the above description, while we other existing measurements such as UL SRS RSRP etc. are also applicable. Also, in our understanding, UE location can be provided not only by LMF but also by UE or PRU UEs as well.  Besides, “no latency requirements” for training could be too early to say. “Relaxed latency requirements” would be more appropriate. Also, existing configuration is meant to perform non AI/ML Multi-RTT and UL-based positioning, whether existing configuration can fulfil the requirements needs further assessment, e.g., it’s possible that AI/ML based positioning is configured as a new positioning method using new IE. In any case, RAN1 input is required to better understand the requirements. |
| Interdigital | Same comment as Q3 |
| vivo | See comments to Q3 and Q4. |
| Xiaomi | As replied in Q1, we understand in all cases, i.e. other than case 2b and 3a, data collection for training and monitoring may be needed if the training and monitoring is done by LMF/gNB.  In general, we are fine with the main bullets. Sub-bullets may need further discussion.  In case 3a, there may be no impact to air interface. |
| NEC | See comment for Q3 |
| LGE | Same as Q3 |

*[Rapporteur to add summary of views]*

1. To be added according to companies’ views…

Continuing with Case 2b, the Rapporteur understands that the following aspects could potentially be considered to design solutions:

1. **The content of the data**
   * Radio measurements:
     + DL-PRS based CIR measurements (reported from UE to LMF over LPP)
   * Non-radio measurements:
     + UE location (ground truth label, obtained at LMF)
2. **The data size** 
   * Data should be collected from multiple UEs, at one- or multiple- UE locations
   * For each UE report for a single UE location
   * For each UE report, the number of bits needed to be reported per UE location will depend on e.g., the type of measurements a UE should report, how the measurement is represented, and how many TRPs are involved
3. **Latency, periodicity, or “efficiency”**
   * There appears to be a relaxed latency requirement for model training
4. **Configuration-related requirements**
   * To study how/whether enhancements are needed

**Q6)** Companies are invited to comment on the above requirements and whether these could be considered for designing data collection solutions for UE-assisted/LMF-based positioning with LMF-sided model?

Note: Companies are encouraged to provide further views or other aspects that should be considered.

|  |  |
| --- | --- |
| **Company** | **Answer / Comments** |
| Apple | See our comments to Q3. For now, we don't think RAN1 details on a) b) c) related to AI/ML based positioning are clear. Thus, RAN2 should wait further RAN1 input. |
| OPPO | The similar view as Q4. |
| Qualcomm | See comment for Q3. |
| Lenovo | Same comment as Q3, and we understand the sub-bullets under a)b)c)d) are given only as examples, since RAN1 is still discussing, e.g., the exact input data for training, and data size. |
| Interdigital | Same comment as Q3 |
| vivo | See comments to Q3 and Q4. |
| Xiaomi | As replied in Q1, we understand in all cases, i.e. other than case 2b and 3a, data collection for training and monitoring may be needed if the training and monitoring is done by LMF/gNB.  In general, we are fine with the main bullets. Sub-bullets may need further discussion. |
| NEC | See comment for Q3 |
| LGE | Same as Q3 |

*[Rapporteur to add summary of views]*

1. To be added according to companies’ views…

### 2.2.3 CSI feedback enhancement

The Rapporteur acknowledges that RAN1 have not provided explicit input to RAN2 concerning requirements for the CSI use cases.

The Rapporteur understands though, that the following could be considered for the initial design of solutions for two-sided CSI feedback compression use case:

1. **The content of the data**
   * Radio measurements:
     + CSI-RS measurements (target H or target precoder for model training/monitoring)
   * Non-radio measurements:
     + Cell ID, area ID, carrier frequency, UE/NW antenna beam configuration/ID, time stamp, UE location, measurement accuracy, measurement resolution, etc. …
   * Monitoring metrics:
     + Intermediate KPIs, e.g., SGCS
     + Legacy CSI based monitoring, additional legacy CSI reporting
2. **The data size** 
   * Data should be collected from UE(s), at one- or multiple-time instances
   * The number of bits needed for reporting a single channel measurement (a target H) depends on how the target channel is represented, i.e., the data format and preprocessing/quantization.
3. **Latency, periodicity, or “efficiency”**
   * Data collection for model monitoring could eventually be a time-sensitive procedure when compared to the model training procedure
4. **Configuration-related requirements**
   * It should be possible to configure a UE to store measurements on multiple occasions and then report the accumulated data to the NW
   * Periodic and event-triggered data collection approaches could be considered for further study
   * Pre-processing may need configuration, depending on how the target channel is represented

**Q7)** Companies are invited to comment on the above requirements and whether these could be considered for designing data collection solutions for spatial-frequency domain CSI compression?

Note: Companies are encouraged to provide further views or other aspects that should be considered.

|  |  |
| --- | --- |
| **Company** | **Answer / Comments** |
| Apple | See our comments to Q3. For now, we don't think RAN1 details on a) b) c) related to AI/ML based CSI enhancement are clear. Thus, RAN2 should wait further RAN1 input. |
| OPPO | The similar view as Q4. |
| Qualcomm | See comment for Q3. |
| Lenovo | Same comment as Q3, and we understand the sub-bullets under a)b)c)d) are given only as examples, since RAN1 is still discussing, e.g., the exact input data for training, and data size. |
| Interdigital | Same comment as Q3 |
| vivo | See comments to Q3 and Q4. |
| Xiaomi | Same as Q3. |
| NEC | See comment for Q3 |
| LGE | Same as Q3 |

*[Rapporteur to add summary of views]*

1. To be added according to companies’ views…

## 2.3 Data collection methods

As discussed online during RAN2#120 and as seen in a few of the companies’ contributions to the previous WG meeting, the MDT framework and mechanisms has been brought up as a candidate.

As per the scope of this email discussion, RAN2 should now focus:

|  |
| --- |
| [...] on to what extent existing methods can be useful including also identifying these existing methods and their potential extensions [...] |

In this regard, the Rapporteur understands that the following frameworks can also potentially be considered:

* MDT,
* UE assistance information (defined in RRC-spec.),
* early idle/inactive measurements,
* RRC measurement reports,
* CSI reporting framework.
* LPP Provide location information

These are all existing procedures/methods that rely on configuration to collect data. It would then be beneficial to understand companies views on this matter. Hence the following question.

**Q8)** Do you agree to consider the above candidate frameworks as starting points to be considered for data collection in this SI?

|  |  |
| --- | --- |
| **Company** | **Answer / Comments** |
| Apple | 1. We have some concern on "early idle/inactive measurement". In current RAN2 identified 3 scenarios, only RRC\_CONNECTED UEs are the target. We are not sure why measurement performed by IDLE/INACTIVE UEs can help. Thus, we suggest to remove it.  2. We are a little confused with the intention of this proposal. Our understanding on the intention of this proposal is:   * RAN2 will analyze and compare these existing framework for data collection. * Based on specific requirements of data collection for AI/ML (discussed in Q3), RAN2 will study whether these existing framework can be reused/extended, or a new framework is required to be introduced.   Maybe rapporteur can confirm whether our understanding is correct. |
| OPPO | General speaking, we’re fine to use all existing data collection frameworks as the baseline, but it will be better if we can go deeper for existing data collection frameworks in this email discussion, i.e. evaluate how it works for each existing data collection framework (Maybe give a general signalling flow for each framework) and also confirm the applied use cases/data types for each framework, that’s one of our main targets for this email discussion, on top of this, we can re-evaluate whether the existing data collection frameworks should be reused or extended based on RAN1 data collection requirements.  In short, it’s hard for companies to make the judgement for any enhancement without aligning the understanding on existing data collection framework as not all delegates working on AI topics are familiar with all existing data collection frameworks. |
| Qualcomm | As previously mentioned, the requirements for data collection for model training, inference, monitoring, updates, etc are different. Therefore, RAN2 needs to evaluate and analyze the existing framework for meeting the data collection requirement for them. Therefore, we have a similar confusion with the intention of this proposal, as Apple. We also believe that intention of this proposal should be:   * RAN2 will analyze and compare these existing frameworks for data collection for training, inference, monitoring, updates, etc one-by-one. * Based on specific requirements of data collection for training, inference, monitoring, updates, etc one-by-one, RAN2 should study whether these existing frameworks can be reused/extended, or whether a new framework is required to be introduced. |
| Lenovo | OK to take legacy data collection frameworks as start point, RAN2 needs to further discuss if/how to make use of them for model training/inference/monitoring/update. Also, for LPP procedure, it would be good to assess the framework taking into account different data source (e.g., UE, LMF, TRP) separately. |
| Interdigital | Agree with the identified frameworks. And as we have indicated above, we think the detailed analysis of these frameworks (features/pros/cons/capabilities/limitations, etc.) should be the main focus of this email discussion. |
| vivo | RAN2 can consider above candidate frameworks as starting points and further evaluate them with regard to how they may satisfy data collection requirement per use case. |
| Xiaomi | We can first evaluate whether the existing functionalities can fulfil the requirement. |
| NEC | We generally agree to consider all the listed frameworks. As a quick our feeling for early Idle/Inactive measurements, we do not assume this is suitable for AI/ML, because this is defined for the specific purpose unlike the others. But ok to discuss. |
| LGE | Agree with the above frameworks for the starting point. |

*[Rapporteur to add summary of views]*

1. To be added according to companies’ views…

If existing methods (as a whole) could not be suitable for the use cases covered in this SI, new approaches might be needed to fulfil the requirements of these.

As there are no clear requirements for the use cases and as per the status of the SI, it is perhaps too early to describe or agree a particular solution (or a set of solutions) that could be used to collect data. However, it would be beneficial to understand whether companies think something is missing from the previous question, or whether there are already some thoughts regarding the procedures/mechanisms that should be used.

**Q9)** Is there any aspect(s) missing in this part that should be addressed by RAN2? Do companies have further views concerning the data collection mechanism(s) that should be used?

|  |  |
| --- | --- |
| **Company** | **Answer/Comments** |
| Apple | Security of data and UE privacy should be considered. |
| OPPO | As suggested in Q8, RAN2 should go deeper for existing data collection frameworks in this email discussion, i.e. evaluate how it works for each existing data collection framework (Maybe give a general signalling flow for each framework) and also confirm the applied use cases/data types for each framework before discussing any enhancement. |
| Qualcomm | We agree with OPPO that RAN2 should go deeper into existing data collection frameworks in this email discussion, i.e. evaluate how it works for each existing data collection framework. Furthermore, this should be done for data collection for training, inference, monitoring, updates, etc one by one instead of bundling them together.. That is RAN2 need to understand different requirements for data collection for different purposes and needs to evaluate the limitation/benefit of existing methods. RAN2 then should study the new framework/method for data collection for different purposes. |
| Lenovo | Similar view as Qualcomm. In addition, we see data collection discussion has strong dependency on the other discussion on model delivery, especially where the AI training functionality locates. Once RAN2 has better understanding on which node will collect data and perform training, RAN2 could have more fruitful discussion on possible enhancements on data collection procedure. |
| Interdigital | We think the detailed analysis of the different existing frameworks listed before Q8 will result in identification of any missing/non-supported features in current frameworks that are needed for the different AI/ML use cases and data collection purposes (once RAN1 has identified the requirements). |

*Rapporteur to add summary of views]*

1. To be added according to companies’ views…
2. To be added according to companies’ views…

## 2.4 RAN1 input and additional comments

Until this point, there has not been any clear collaboration between RAN1 and RAN2. This has created some confusion regarding the WG split and, arguably, it has also slowed-down RAN2 progress.

In this regard, an option would be for RAN2 to directly ask RAN1 for further information that could help RAN2 in the design of solutions for this SI. Alternatively, RAN2 could continue the work within their scope and inform RAN1 of concerning RAN2 agreements.

**Q10)** What approach do you think RAN2 should adopt concerning the collaboration with RAN1? Do you think RAN2 should ask for RAN1 input in particular areas? Do you think RAN2 should simply keep progressing and inform of concerning agreements to RAN1 when necessary? Should another or no different approach as the one followed now should be taken?

|  |  |
| --- | --- |
| **Company** | **Answer/Comments** |
| Apple | We prefer "RAN2 should simply keep progressing and inform of concerning agreements to RAN1 when necessary". Specifically for data collection:   * Our understanding is RAN1 need more work to discuss and define requirement. And RAN2 should wait RAN1's progress. Even if RAN2 send LS to RAN1, we don't think it can speed up RAN1's work. * Of course, if RAN2 can identify some specific questions for RAN1 to speed up progress, we are open to it. But for now, we don't identify any useful questions. |
| OPPO | We slightly prefer ‘*RAN2 could continue the work within their scope and inform RAN1 of concerning RAN2 agreements*’ for data collection discussion, but as suggested in Q8, RAN2 should first do the alignment on each existing data collection framework before discussing any enhancement, which is necessary for further discussion on data collection. As for the new data collection requirements, RAN2 should just follow RAN1 guidance if only RAN1-led use cases are considered. |
| Qualcomm | Our understanding is that for certain things like,   * The content of the data * The data size * Latency, periodicity, or “efficiency” * Configuration-related requirements   RAN2 should wait for RAN1 progress as RAN1 is determining and evaluating them.  RAN2 should start discussing whether the existing framework is sufficient for different purposes (training, inference, monitoring, updates, etc). We believe that RAN2 should study data collection requirements and require a framework (existing/new) for training, inference, monitoring, updates, etc one by one instead of bundling them together. |
| Lenovo | “*RAN2 could continue the work within their scope and inform RAN1 of concerning RAN2 agreements*”. RAN2 can make qualitative analysis and try to reach some common understanding, e.g., requirements for AI training/inference/monitoring/update could be different, assess legacy framework, and identify possible enhancements.  To determine which enhancements are necessary, RAN2 would need more quantitative analysis and shall wait for RAN1 progress w.r.t data size and latency requirements.  For the content of the data, RAN2 can simply follow RAN1 conclusion. |
| Interdigital | One approach could be for RAN2 to inform RAN1 about the capabilities/limitation of current data collection frameworks, after the detailed analysis is undertaken. |
| vivo | We prefer "RAN2 should simply keep progressing and inform of concerning agreements to RAN1 when necessary". |
| Xiaomi | We understand RAN1 was considering to provide some info by LS, but failed to reach consensus. So, maybe we can first ask some question about key requirements. For example, the size of AIML model and latency of AIML model transmission. |
| LGE | Agree with QCT |

**Q11)** Is there anything else companies would like to bring up concerning this email discussion? (e.g., discuss additional topics)

|  |  |
| --- | --- |
| **Company** | **Answer/Comments** |
| OPPO | Evaluate how it works for each existing data collection framework (Maybe give a general signalling flow for each framework) and also confirm the applied use cases/data types for each framework. |
|  |  |
|  |  |
|  |  |
|  |  |

*[Rapporteur to add summary of views]*

1. To be added according to companies’ views…

# 3 Conclusion

In the previous sections we made the following observations:

[Observation 1 RAN2 should not rely on RAN3 regarding architecture or functionality-to-entity mapping matters, since RAN3 does not have TUs assigned for this SI.](#_Toc122071371)

[Observation 2 For UE-sided AIML models, UEs do not necessarily need to report the collected data for model training to the gNB/LMF.](#_Toc122071372)

[Observation 3 RAN1 can later provide requirements (e.g., related to configuration, signalling, etc.) which could facilitate the process of UE data collection for UE-sided AIML models.](#_Toc122071373)

[Observation 4 There seems to be a need for RAN2 to start by focusing on gNB/LMF-sided AIML models. Since for these cases, the UE needs to be configured to report collected data (e.g., measurements, performance metrics) to the gNB/LMF.](#_Toc122071374)

[Observation 5 To be added according to companies’ views…](#_Toc122071375)

Based on the discussion in the previous sections we propose the following:

[Proposal 1 To be added according to companies’ views…](#_Toc122071362)

[Proposal 2 To be added according to companies’ views…](#_Toc122071363)

[Proposal 3 To be added according to companies’ views…](#_Toc122071364)

[Proposal 4 To be added according to companies’ views…](#_Toc122071365)

[Proposal 5 To be added according to companies’ views…](#_Toc122071366)

[Proposal 6 To be added according to companies’ views…](#_Toc122071367)

[Proposal 7 To be added according to companies’ views…](#_Toc122071368)

[Proposal 8 To be added according to companies’ views…](#_Toc122071369)

[Proposal 9 To be added according to companies’ views…](#_Toc122071370)

# 4 References

1. [RP-213599](http://www.3gpp.org/ftp//tsg_ran/TSG_RAN/TSGR_94e/Docs//RP-213599.zip), “New SI: Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR Air Interface”, TSG RAN, RAN#94-e, Dec 2021
2. [RP-221348](http://www.3gpp.org/ftp//tsg_ran/TSG_RAN/TSGR_96/Docs//RP-221348.zip), “Revised SID: Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR Air Interface”, TSG RAN, RAN#96, Budapest, Hungary, June 2022
3. [R1-2205695](http://www.3gpp.org/ftp//tsg_ran/WG1_RL1/TSGR1_109-e/Docs//R1-2205695.zip), “Session notes for 9.2 (Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR air interface)”, Ad-hoc Chair (CMCC), RAN1#109-e, May 2022
4. [R1-2208145](http://www.3gpp.org/ftp//tsg_ran/WG1_RL1/TSGR1_110/Docs//R1-2208145.zip), “Session notes for 9.2 (Study on Artificial Intelligence (AI)/Machine Learning (ML) for NR air interface)”, Ad-hoc Chair (CMCC), RAN1#110, Toulouse, August 2022

1. [R2-2210677](http://www.3gpp.org/ftp//tsg_ran/WG2_RL2/TSGR2_119bis-e/Docs//R2-2210677.zip), “RAN2 Work Plan for Rel-18 SI on AI/ML for NR air interface”, Ericsson, Qualcomm Inc., RAN2#119-e, Oct. 2022
2. [R1-2210690](http://www.3gpp.org/ftp//tsg_ran/WG1_RL1/TSGR1_110b-e/Docs//R1-2210690.zip), “Session notes for 9.2 (Study on AI/ ML for NR air interface)”, CMCC, RAN1#110bis-e, Oct. 2022
3. [R1-2212845](http://www.3gpp.org/ftp//tsg_ran/WG1_RL1/TSGR1_111/Docs//R1-2212845.zip), “Session notes for 9.2 (Study on AI/ ML for NR air interface)”, Ad-hoc Chair (CMCC), RAN1#1111, Toulouse, November 2022