The ATIS IIF response is provided in red.

Dear Dan O'Callaghan,

Thank you for the additional information on the approach ATIS-IIF is taking to improve response time for objective video quality models. The value of this vision is obvious – a process with fast turnaround to promote new techniques and services. The approach taken in the past by VQEG and the ITU is thorough but unfortunately slow. SG9 is interested to see how this vision progresses, as details are added.

We understand that this is a work in progress, and recognize that many issues mentioned herein have likely already been discussed by ATIS-IIF. This liaison dwells on potential pitfalls only as this appears to be the area that our assistance might be most helpful: to encourage discussion on issues that perhaps you have not considered, or to add details to these documents further specifying these issues.

We identified seven potential problems with the approach of using secret database to validate models, as described in the documents forwarded to us.

1. **Qualification of ITL:** Who oversees ITLs and how? How will you ensure that ITLs keep the database secret? Who controls the ITL membership? How would an ITL resign without damaging the secrecy of the database? Who legally owns the database? If a part of the database is owned by one company, what happens if that company resigns from ITL? If PVSs (processed video sequences) are owned by a company that wishes to resign, it may not be possible for that company to transfer the PVSs to a remaining ITL.

   All of these matters are governed by a third party organization. ATIS IIF may or may not be that third party organization. The third party organization may validate the credibility of ITLs. There will be individual agreements between ITLs and third party organizations, and the agreements will ensure the integrity and confidentiality of the databases. There may not be a single company owning the entire database. The databases may be owned by ITLs and/or the third party organization. Licenses should be written to keep the database secret.

2. **Funding/Fee:** Who will decide the fee schedule, and what variables will the fee schedule depend upon?

   Generally, the model developers pay for most services. The actual fee schedule will be negotiated between the third party organization, the ITLs and model developers. There are two generic models:

   - **The ITL and model developers negotiate fees**

   - **The third party organization develops the fee schedule**

3. **Size of Database:** Once a set of sequences have been used to validate a model, their value as validation data will be reduced. That is, every piece of information given to model developer and model user reveals information about the database. What constraints are placed to ensure the continued viability of the database?

   The intention is to have a database that is sufficiently large and varied to make tuning difficult. Furthermore, the ITL will randomize the selected sequences used for individual model validations.

4. **Frequency of Submission:** If developers are allowed to submit models many times, they can optimize their models to the test data. The model developer could, as an extreme example, submit 100 models each of which is actually a single parameter, then train a model using the five parameters that worked best.

Based on the use of randomized databases mentioned above, we believe any practical instance of this kind of optimization should not be a significant concern. Since model developers pay for each validation, and an ITL will only be able to do so many validations within a given time period, there are also practical and financial constraints to the number of possible submissions.

5. **Detailed Report:** It appears that the Detailed Report gives too much information to the model developer, such that the database is no longer entirely secret after only one use. If MOS and predicted MOS values are made available to developers, as indicated in the Detailed Report, then it is possible for the developers to reverse-engineer.

We believe the randomization principle mentioned above will minimize any significant ability to understand the database in detail thus reducing this concern significantly. The report will not reveal the content of the PVS, so that the MOS score cannot be traced back to a specific PVS.

6. **Small Companies:** This approach may give advantages to companies with large financial resources since they can afford to regularly submit new models.

We believe a reasonably structured fee schedule will promote involvement by smaller companies and reduce the practical impact of your concern. We hope this streamlined process will lower barriers to entry and promote participation.

7. **Authentication of the Validation Process:** Since the database is secret, it is not clear how to demonstrate the validation tests are balanced and represents all test conditions to the developers and potential users.

Information about the source content and test conditions will be available publicly, as specified in Section 7.6 of the document. Furthermore, the professionalism of the third party organization and ITLs is assumed to be at a high level.

It appears that the following ideas might be used to improve the draft validation process:

1. **Large Pool for Each Category:** The specifications should indicate a minimum size of the database for each individual category. This specification should identify the minimum number of unique source scenes, and the minimum number of different systems tested, as well as the minimum coder settings, decoder settings, and network settings examined.

We have edited WT-066 to provide some general guidelines on the size of the database and numbers of SRCs for each genre.

2. **New Data:** Potential issues with model developers training on the database can be addressed by requiring that new data be added to the database each year. The issue of how to anchor the new MOS with the existing MOS will need to be addressed.

We anticipate that the database will grow and ITLs will continue to grow the databases. We expect this to be done so that any sequences added to the database will not be significantly different from a regular randomized selection. The report will also clearly identify the database version or year.

3. **Limit Detailed Report:** To avoid revealing excessive information on the database to users and model developers, the Detailed Report should not contain specific information on any individual PVSa.

See response to item 5 above regarding Detailed Report.

4. **ITL Report:** The ITL should issue a report each year. This report should not include company names, but rather a summary of the ITL's activites (e.g., number of models tested, summary of outcomes, size of the database, summary of conditions in the database, identity of every company involved in the ITL).

We appreciate this suggestion. The third party should issue a report each year and this report can list the number of models tested, summary of outcomes, size of the database, summary of conditions in the database and identities of every ITL.

5. **Maximum Model Submission:** Establish a maximum number of models that can be submitted from any one company each year.

The number of tests done on each model is recorded in the summary report and we have now included a limit in WT-066.

Regarding SG9 members' interest in becoming ITL, this will likely depend upon decisions that ATIS-IIF makes regarding fees, legal agreements and final details of the plan. SG9 will continue to follow the development of the ATIS-IIF validation procedure with interest.

We believe a collaborative process would be most appropriate.

We look forward continued coordination on this issue and hope that we can be mutual assistance in this area of video quality assessment.

_____