 ISO/IEC JTC 1/SC 29/WG 8 N0041

**ISO/IEC JTC 1/SC 29/WG 8**

**MPEG GENOMIC CODING   
Convenorship: SNV (SWITZERLAND)**

**Document type:** Output Document

**Title:** Evaluation Procedure for the Joint Call for Proposals on New Advanced Features and New Technologies for MPEG-G

**Status:** Approved

**Date of document:** 2021-05-1

**Source:** ISO/IEC JTC 1/SC 29/WG 8

# Expected action: None

**Action due date:** None

**No. of pages:** 5 (with cover page)

**Email of Convenor:** marco.mattavelli@epfl.ch

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg8>

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 8**

**MPEG GENOMIC CODING**

**ISO/IEC JTC 1/SC 29/WG 8 N0041**

**Online – April 2021**

|  |  |
| --- | --- |
| **Title** | **Evaluation Procedure for the Joint Call for Proposals on New Advanced Features and New Technologies for MPEG-G** |
| **Source** | **WG 8, MPEG GENOMIC CODING** |
| **Serial Number** | **20376** |

# Introduction

The Evaluation Procedure described in this document will be used to assess proposals in response to the “Joint Call for Proposals for New Advanced Features and New Technologies for MPEG-G.”

This assessment will identify both the proposal which will become part of the Test Model 0 (TM0) and other proposals from which a set of technologies will be possibly added to or integrated with the TM0. The addition and integration will be achieved through Core Experiments that will be specified by dedicated MPEG documents. A timetable and the procedure for the verification process are specified in the Call for Proposals.

The objective is to define a new profile for the ISO/IEC 23092 standard series. It should be noted that the procedure described above implies that proposals will be evaluated whether they fulfill requirements entirely or partially, but also in the case that they provide technology covering only a subset of the requirements.

This Evaluation Procedure will be used by:

* Proponents of technology/solutions for extensions and improvements of ISO/IEC 23092 series, both in answering to the Call for Proposals, and in self-assessing their proposal before submitting it;
* MPEG, to validate the self-assessment of the proponents and proceed to the evaluation of the proposed technologies.

The requirements defined in the Call for Proposals and Requirements are divided into major needs (specified by a “shall” in the requirement text) and objectives (specified by a “should” in the requirement text). Major requirements shall be met by TM0 while objectives are less stringent even if highly desirable.

# Encoded bitstreams and test data

Encoded bitstreams demonstrating the value, performance and satisfaction of the proposed technology shall use the test material specified in (N0024) “MPEG-G Genomic Information Database”

# Evaluation Procedure

This section describes the test cases that will be performed to validate the fulfillment of the requirements listed in the Call for Proposals [1] and Requirements [2]

## General functionality

Every submission shall clearly state:

1. which data types are supported by proposed technology
2. which performance the proposed technology is providing versus current technology specified in ISO/IEC 23092
3. which new functionality the proposed technology is providing

## Third Generation Sequencing (Long Read) Coding Mode

This section describes the process to assess the coding efficiency of submitted solutions for third generation sequencing coding mode.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test case ID** | **Input Items** | **Test case steps** | **Reqs** |
| Indicate the ID of the input file used to demonstrate results | Any input file containing genomic data in the MPEG-G repository as detailed in N0024 “MPEG-G Genomic Information Database” and all updates of this document issued at the 135th and 136th MPEG meetings. | 1. Compress data with new technology 2. compare encoding, decoding performance 3. compare compression rates with current standard ISO/IEC 23092 technology and with the most recent version of the “MPEG-G Performance Benchmark” | 4.1.1 to 4.1.9 |

## Machine Learning Coding Mode

This section describes the process to assess the coding efficiency of submitted solutions for machine learning coding mode.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test case ID** | **Input Items** | **Test case steps** | **Reqs** |
| Indicate the ID of the input file used to demonstrate results | Any input file containing genomic data in the MPEG-G repository as detailed in N0024 “MPEG-G Genomic Information Database” and all updates of this document issued at the 135th and 136th MPEG meetings | 1. Compress data with new technology 2. compare encoding, decoding performance 3. compare compression rates with current standard ISO/IEC 23092 technology and with the most recent version of the “MPEG-G Performance Benchmark” 4. provide results of the cost of representing and storing neural network models and specifications | 4.2.1 to 4.2.11 |

## Improved Compression of Genome Sequences, Quality Scores and Metadata

This section describes the process to assess the coding efficiency of submitted solutions for improved compression of genome sequences, quality scores and metadata.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test case ID** | **Input Items** | **Test case steps** | **Reqs** |
| Indicate the ID of the input file used to demonstrate results | Any input file containing genomic data in the MPEG-G repository as detailed in N0024 “MPEG-G Genomic Information Database” and all updates of this document issued at the 135th and 136th MPEG meetings | 1. Compress data with new technology 2. compare encoding, decoding performance 3. compare compression rates with current standard ISO/IEC 23092 technology and with the most recent version of the “MPEG-G Performance Benchmark” | 4.1.1 to 4.1.9 |

## Graph Reference Genome Support

This section describes the process to assess the coding efficiency of submitted solutions for graph reference genome support.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test case ID** | **Input Items** | **Test case steps** | **Reqs** |
| Indicate the ID of the input file used to demonstrate results | Any input file containing genomic data in the MPEG-G repository as detailed in N0024 “MPEG-G Genomic Information Database” and all updates of this document issued at the 135th and 136th MPEG meetings | 1. Compress graph reference genome with new technology 2. compare encoding, decoding performance 3. compare compression rates with current standard ISO/IEC 23092 technology 4. demonstrate how reads aligned to the new representation of the graph reference genome can be used with existing ISO/IEC 23092 compression technologies | 4.3.1 to 4.3.10 |

## Support for Interchange of Genomic Information with Other Standards

This section describes the process to assess the coding efficiency of submitted solutions supporting the interchange of genomic information with other standards.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test case ID** | **Input Items** | **Test case steps** | **Reqs** |
| Data can be either provided by responder or inserted into MPEG-G Database with associated ID | FHIR Genomic Reporting Implementation Guide Resource data extracted from an EHR to be included into MPEG-G metatdata and MPEG-G data and metadata extracted from an MPEG-G file to be used to fill and EHR record via FHIR Resources | 1. Demonstrate how data can be converted between FHIR and MPEG-G retaining the full information content 2. compare input and output data 3. assess search and browsing capability 4. assess protection and selective access capabilities | 4.4.1 to 4.4.3 |

# Selection criteria

Criteria to rank technologies overall will be:

1. New requirements coverage versus current ISO/IEC 23092 technology
2. New functionality coverage versus current ISO/IEC 23092 technology
3. Performance improvement versus current ISO/IEC 23092 technology performance
4. Usage of and compatibility with relevant standards

The proponents are required to characterize the solution on the points above and provide evidence (code, examples, statistics, etc.) of the results.

*Nota bene*: additional measured performance such as:

1. computational complexity
2. processing time
3. memory requirements

might be used to further rank proposals.

# References

|  |  |
| --- | --- |
| [1] | ISO/IEC JTC 1/SC 29/WG 8, "N00xx - Final Joint Call for Proposals on New Advanced Features and New Technologies for MPEG-G," On-Line meeting, April, 2021. |
| [2] | ISO/IEC JTC 1/SC 29/WG 8, “N00xx - Requirements for Joint Call for Proposals for New Advanced Features and New Technologies for MPEG-G," On-Line meeting, April 2021. |