INTERNATIONAL ORGANISATION FOR STANDARDISATION
ORGANISATION INTERNATIONALE DE NORMALISATION
ISO/IEC JTC 1/SC 29/WG 11
CODING OF MOVING PICTURES AND AUDIO

ISO/IEC JTC 1/SC 29/WG 11
MPEG2017/m41073
July 2017, Torino, Italy

Source: **DMAG-UPC**
Status: **Proposal**
Title: **Use case for referencing external existing genomic information from MPEG-G**
Authors: **Jaime Delgado, Daniel Naro, Silvia Llorente (Distributed Multimedia Applications Group – Universitat Politècnica de Catalunya)**

Table of Contents

1 Purpose

This document describes a possible new use case where genomic information is stored in external (possibly existing) files referenced from MPEG-G files.

2 Proposal of new use case

2.1 Reference external existing genomic information

UCId	GI-UC08
Use Case Name	Reference external existing genomic information
Description	Currently, there is an important amount of genomic information represented in formats like SAM, BAM, FASTQ, etc. When starting to use MPEG-G, research centres storing this information may want to maintain their current files, without transcoding them to MPEG-G. In order to have some of the advantages of MPEG-G, like the use of metadata, privacy rules or API operations, we could generate only the headers of MPEG-G file format without including the genomic data but references to existing external files. In this way, some search and protection features may be immediately applied, without involving extra space nor processing time.
Actors	Analyst, tool to associate existing file format to MPEG-G file, Organisations storing genomic information
Assumptions	The genomic information is conveniently stored and accessible from the header files
Actions	1. The analyst generates a new empty MPEG-G file. 2. The analyst associates existing genomic information files to the MPEG-G file, following the defined structure (Dataset group, Dataset, etc.). 3. The analysts saves the MPEG-G file containing headers plus references to the external genomic information representation files.
Inclusion or Extensions	The generated MPEG-G file could be a combination of referenced files together with native MPEG-G represented information. In this case, the file used in step 1 does not need to be an empty file, but it could be any MPEG-G file.
Issues	In case an existing MPEG-G file is used, the memory and disk management of the file could be much more complex than when only references are considered, due to the amount of information it may contain.