

**INTERNATIONAL ORGANISATION FOR STANDARDISATION  
ORGANISATION INTERNATIONALE DE NORMALISATION  
ISO/IEC JTC1/SC29/WG11  
CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG2017/M42590  
April 2018, La Jolla, US**

**Source: DMAG-UPC**  
**Status: Proposal**  
**Title: Comments on Draft Text of ISO/IEC DIS 23092-1 Transport and Storage of Genomic Information**  
**Authors: Daniel Naro, Jaime Delgado (Distributed Multimedia Applications Group – Universitat Politècnica de Catalunya)**

**Table of Contents**

1	Introduction	2
2	Comments	2
3	Comments related to other parts	2
4	References	3

## 1 Introduction

At the 121<sup>st</sup> MPEG meeting, a Draft DIS of MPEG-G part 1 was produced [1]. This contribution provides a few comments on that document and an additional related comment.

## 2 Comments

- In Table 4: clarify that block is not a `gen_info` element (maybe with a different font).
- In Table 12: change “`gen_info`” with “`lbl`”.
- Issue with `dscn`: currently, it is not possible to know where the blocks end, as there might be a protection box (the sentence saying that blocks’ size equal to `dscn` size – `dshd` size has to be changed). The issue can be corrected by changing the position of all boxes to be prior to the blocks, or adding information on the total size of the blocks.
- Random access / representation of the file structure in memory would be simplified if Access unit header had always information about the sequence the information is mapped to (i.e. also when `MIT_flag == 1`). This would remove the need to test the `AU_byte_offset` to determine which AU is being indexed.
- In last sentence of 6.5.2.1 introduction: the special value is not only used for `block_byte_offset`, but also to indicate that the access unit is empty (see, for example, `AU_START_POSITION`). Maybe consider not having empty AUs.
- Issue with the use of “Descriptor Id” (for example in Dataset header or MIT): `desc_ID` as used in Dataset Header refers to the index in the list of available descriptors, not to the actual value of the descriptor id. This could be clarified by changing to `desc_index`, for example. Similar issue in MIT: `desc_ID` refers to the index in the list of available descriptors, but could be understood as the ID of the descriptor identifier as specified in Table 10 of ISO/IEC 23092-2.
- Stream container: remove metadata.
- Descriptor stream header: remove protection flag, which is outdated (it is a legacy from the GENIFF proposal).
- Table 9- Reference box syntax defines `seq_count` to be of type 32 bits, but in table 17 the `seq_count` is of type 16 bits: a dataset header could not use all the defined sequences. Consider which size is more appropriate and update all occurrences.

## 3 Comments related to other parts

- There are multiple names used for the concept of the access unit type (`au_type`, `class_id`, `class_type`). Consider which name is preferred and update all occurrences.

- Part 1 defines the position in the genome to be 32 bits (Table 29 au\_\*\_position, extended\_AU\_\*\_position, and same in auhd), but the mpegg\_record defined in table 64 of part 2 defines the mapping position to be of type 64 bits. Consider which size is more appropriate and update all occurrences.

## **4 References**

[1] ISO/IEC JTC 1/SC 29/WG 11, N17320 “Draft Text of ISO/IEC DIS 23092-1 Transport and Storage of Genomic Information” Gwangju, January 2018