

UMID – Unique Material Identifier

A Unique Material Identifier (UMID) is a special code that is used to identify audiovisual (AV) materials. The UMID is a core component in tagging AV content to enable its reliable access and tracking, especially in networked storage, production and dissemination systems. Its purpose is to provide unambiguous identification of material through the production and emission chain, as well as to make possible the reliable linking of essence with its metadata. The UMID is a locally generated and globally unique identifier, standardised by the Society of Motion Picture and Television Engineers (SMPTE), and presents a key component in digital media asset management systems.

Description

The UMID is a specialist identifier used in digital audiovisual systems. It should be automatically created with the essence and remain associated with (parts of) a content object throughout its lifetime. The main purpose of UMID is to identify and locate material in storage systems, to provide constant identification throughout production, archival storage, transmission and dissemination, and to create links between metadata and essence. It allows the management of program copies and even unfinished materials. UMID has played a key role in the success of open-source file interchange formats like MXF (Material Exchange Format) /AAF(Advanced Authoring Format) and is today supported by leading manufacturers in the AV industry.

The UMID is a low-level system component, defined as class 1 metadata within the SMPTE metadata dictionary (referring to: identification and location, local /global scopes), used for identifying a particular instantiation of specific essence. Such a code is usually randomly machine generated (e.g. a random number algorithm combined with a (registered) identifier, in the case of UMID following the SMPTE format specification), and is not specifically designed to be human readable. Therefore its use is expedient in automated systems and may be invisible for users.

The UMID differs from many unique identifiers in that the number does not necessarily depend on a pre-registration process, but can be generated at the point of material origination without reference to a central resource. A key property of a UMID is that it is also possible to use it simply as a globally unique dumb number (i.e. an opaque string of no intrinsic meaning). The UMID itself does not directly provide identification of copyright, nor does it identify program content or works. It acts as a look up ID number (e.g. a pointer via a database) for metadata rather than interleaving the metadata directly with the essence.

What exactly does UMID identify?

The Basic UMID uniquely identifies either:

- a single instance of a clip of material, or
- a bounded group of instances of clips of related material.

The Extended UMID allows for additional metadata information (the “source pack”) to be added.

The UMID offers a single frame granularity, so that it can be used to identify e.g. a single frame of video or audio, a still image file, a related sequence of some still images in a single file, a continuous audio or video, up to an editorially-complete program item (comprising video, picture, multiple-track audio and data elements) or a complete program in a contiguous sequence.

UMID Format Specification

In 2000, SMPTE released the international standard SMPTE 330M, requiring a UMID to be assigned to all digital audiovisual files and sub-files. The standard was updated in 2004 as SMPTE 330M-2004, aiming to be backwards compatible.

The data representation of the UMID is defined in form of a dual approach:

- 1) 32 Byte (mandatory) Basic UMID with the minimal components necessary for unique identification (the essential metadata), containing a universal label (a registered ISO label, administered by SMPTE), followed by a specification of the nature of the audiovisual material, length description (defines the length of the following byte stream) and a material number (identifies different representations of materials, e.g. copy or version of a clip)

The basic UMID can be extended with another 32 bytes of content descriptive information to create an Extended UMID (EUMID):

- 2) 32 Byte (optional) of metadata source pack, containing information on the creation time and date, geospatial coordinates of the recording location and country/organization/user information (the signature metadata).

Together these basic and extended UMIDs have a defined length of 64 bytes:

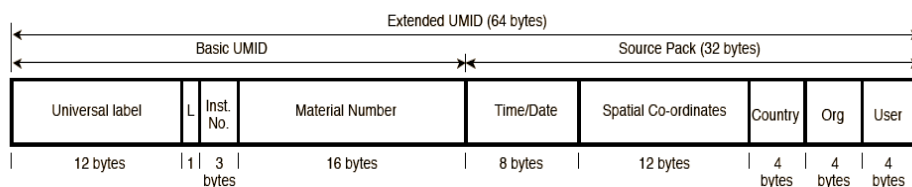


Fig.1: Basic and extended UMID Structures

Applications

Within a networked system, multiple identical or near-identical copies of the content may, and they must be referred to, independently of their location. It is also necessary to refer to content which has been temporarily removed from the system, or to refer to original sources even before the final content has been created. Even in a server-only-based solution, separation of metadata and essence is attractive from the viewpoint of access speed. The European Broadcasting Union (EBU) requires the UMID to be implemented in all systems that create, process or store program material. Basic UMID is mandatory for all applications and Extended UMID is optional.

• Media asset management systems

Once a search through an online database of audiovisual metadata has retrieved a certain number of UMIDs, these will then enable access by cross-referencing between data belonging together. UMIDs enable the application of container formats and complex object-structured wrapper formats (e.g. MXF/AAF).

• Field acquisition

The UMID generated automatically during field capturing may be directly transferred to a laptop communicating with the recording device. Consequently, the UMID will ensure that any additional metadata stored on the laptop during fieldwork will be unambiguously linked to the audiovisual material.

• Easy material retrieval through efficient database searching

Thanks to its Time/Date and Spatial components (assigned on the frame level) the Extended UMID is ideal when it comes to searching for certain scenes in a database of metadata or in the audiovisual material itself.

Some practical considerations: Updating UMID

At every operational step in the production chain – or, more generally: once a new file, processed from a previously existing file, has to be stored – the UMID needs to be updated to guarantee identification of the materials. Two ways are possible: by assigning a completely new UMID to the new file (including new material number and zero instance number), or by merely changing the instance number. The former is a guarantee that the UMID will indeed preserve its global uniqueness in future operations and is the default rule recommended by the EBU. The latter will maintain obvious association with the original data (through the same material number) and may thus be preferred in cases where different formats have to be generated from one and the same original file (e.g. video coding, synchronised audio components of one source etc.). The original source may also still be clearly identified if only the information of the Basic UMID of the new file is altered, with the additional data of the Extended UMID remaining unchanged.

Pros and cons

The merits of machine readable (dumb) identifiers compared to expressive unique identifiers are frequently discussed. Dumb identifiers are quickly applied and do not require human intervention. Their generation mechanism guarantees uniqueness, and most systems are able to allocate a dumb identifier the moment data are saved.

However, the randomness and arbitrariness of dumb identifiers means that machine- and software dependent applications are necessary to show how the different files generated in the life-cycle of a digital resource connect. Comparatively sophisticated database software is required to be operational at many points in a production and dissemination chain. For most non automated applications therefore, a better solution is the use of human readable expressive identifiers.

Further information and resources:

The SMPTE 330M and 330M-2004 standards are available from <http://www.smpte.org/>

SMPTE RP 205-2000 - Application of Unique Material Identifiers in Production and Broadcast Environments. Available from <http://www.smpte.org/>

The UMID registration is administered by <http://www.smpte-ra.org/>

SMPTE metadata dictionary <http://www.smpte-ra.org/mdd/>

EBU Technical Recommendation R108-2001: "High-level rules for systems implementing the SMPTE UMID" http://www.ebu.ch/CMSimages/en/tec_text_r108-2001_tcm6-4697.pdf

EBU Technical Statement D92-2001: "System Requirements for the unique identification of material in broadcast production" http://www.ebu.ch/CMSimages/en/tec_text_d92-2001_tcm6-4721.pdf