



## A Quick Tour of Wrappers and MXF

Audio/Video files are the golden assets of a broadcast facility and production company. There are 100's of different media file formats for video, audio, graphics, text, metadata and more. It's a potpourri enough to confuse anyone. However, at a high level there are just two types of media file formats; *essence-only*<sup>1</sup> and *wrapper-based*. A wrapper is sometimes called a container.

Here are few file formats you may have heard of; MPEG2, JPG, HDV, DV, and MP3. These are *essence-only* formats. On the other hand, MOV (QuickTime), AVI, and MXF are *wrapper* formats. How do they differ? Let's see.

As a paper bag can carry different types of sandwiches and chips, a wrapper file can, in principal, carry various types of A/V essence and other data. Wrapper formats apply mapping recipes to pack the media data inside. Roughly, a wrapper is similar to the common ZIP file that carries many different formats simultaneously in one package. Some wrapper formats only support a few essence types while others support a wide range.

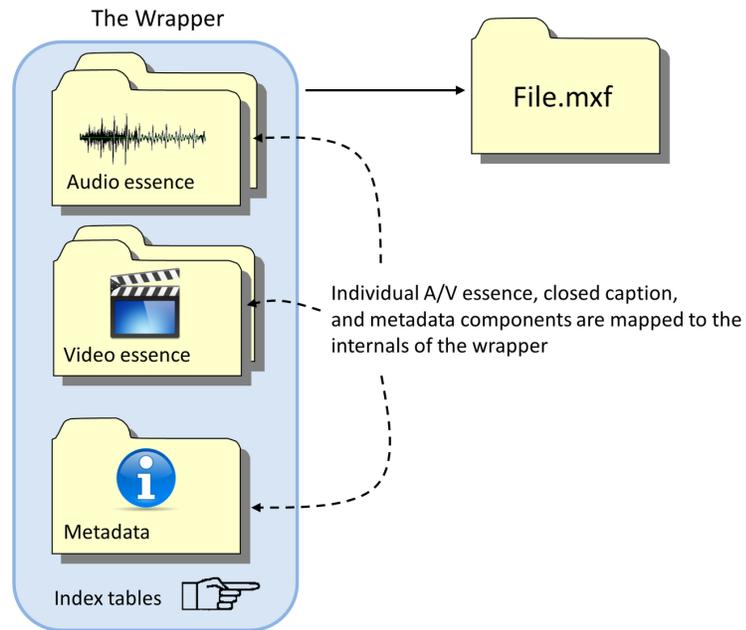
Importantly, a wrapper is not a pure compression format like MPEG video or MP3 audio! This is a common confusion. Wrappers may carry compressed data but this does not make them a codec format.

Why use wrappers for media? The wrapper is an ideal way to guarantee that packaged A/V, metadata (ex, slate), closed caption text and critical index points stay synchronized. Managing time is central to the file's usability. Wrappers can carry many different kinds of essence and keep them aligned; this is non-trivial. Plus, the inside pieces won't become lost -- they are bound together by the wrapper. The wrapper format also defines how metadata should be formatted to guarantee wide use and interchange.

The figure over illustrates the wrapper concept. The internals are composed of time-interleaved (or not) A/V streams and other data types. A common use case is to pack one video program with 4 stereo (8 channels total) foreign language audio tracks. Additionally, there may be a closed caption text stream for each language.

---

<sup>1</sup> An essence-only format is defined as basic A/V or graphic data, compressed or not. For example, JPEG data is graphic essence and PCM audio is raw sound essence. Metadata and closed caption subtitles, on the other hand, are not strictly essence but are fundamental data types.



The Essentials of a Wrapper File

So, is there an “ideal wrapper format” for the media facility? Well, many have been tried over the years but the Material eXchange Format (MXF) has been custom designed for use by the media professional. Its characteristics make it ideal for the capture, editing, distribution, playout and archiving of media. In 2012, cameras export it, editors compile with it, facilities transfer it, and cinemas show movies using it. Today, there are 100’s of millions of MXF files in use. MXF is the true lingua franca of the media facility.

Its life began in 1999 and initial standardization was started by SMPTE in 2003 with essence mapping rules and revisions added along the way. SMPTE, with industry advocates, has standardized essence mapping rules<sup>2</sup> for MPEG2, H.264, JPEG2K, DV, VC-1, VC-2, VC-3,

audio formats, metadata and closed caption formats. MXF is truly a rich format for the media facility.

A key MXF attribute is flexibility but this has become the format’s greatest liability. The universal nature of MXF permits too many ways to get the same job done leading to some incompatible implementations. Flexibility is a double edged sword for achieving interoperability. This is where the AMWA specs step in; reduce MXF’s complexity by limiting what it can do. AMWA’s Application Specification family is a tonic for MXF, creating solid usage templates that provide *Interoperability for Real*.

Contributed by AI Kovalick  
Media Systems Consulting 1/2012

<sup>2</sup> MXF supports more essence mappings than listed here. See SMPTE ST 377 the overall master standard to dig deeper.