

# Preserving Moving Images

## Data Types Series

Artefactual Systems and the Digital  
Preservation Coalition



## DPC Technology Watch Guidance Note

July 2021



Digital Preservation Coalition

## The Data Type Guidance Note Series

Each Guidance Note in the Data Types series is designed to provide a primer on the current state of community knowledge about data types commonly encountered by those seeking to preserve digital holdings. Digital preservation is about keeping information findable, usable, and trustworthy over the long-term. The best approach for any repository will vary according to the scope and content of its holdings, available resources, and the expectations of its funders and users. There are however, broadly applicable good practices that have been established as a result of many years of research, practical implementation, and consensus building. These are presented here as a starting point, along with additional resources for further exploration.

This series of Data Type Guidance Notes has been authored by staff at Artefactual Systems in collaboration with the Digital Preservation Coalition. These notes have been developed in conjunction with the UK Nuclear Decommissioning Authority.

Digital preservation is an evolving field and continues to change and develop in response to external drivers and fresh challenges. New formats, standards, and examples of good practice will emerge over time and the information contained within this report will need to be updated. We welcome comments and feedback to: [info@dpconline.org](mailto:info@dpconline.org).

## 1 Overview

'Moving image' is a term used to describe visual works that have the appearance of movement ([SAA, 2020](#)). This medium has been implemented using a wide range of technical innovations and practices, beginning with the advent of film production in the 1890s, continuing with analogue video formats stored on magnetic tape and optical media stored on discs, to present day works using streaming media, 360° video, Virtual Reality, and Augmented Reality technologies.

Digital moving images can be structured in various ways, including:

- encoded data streams (e.g. H.264) that are wrapped in format containers (e.g. Matroska Multimedia Container).
- a series of raster image files where each represents one frame of film (often from digitization of analogue film reels), typically wrapped in a container and accompanied by contextual metadata (e.g. DPX).
- data that is exposed to the viewer only as a stream available through a web or phone-based app (e.g. TikTok or Instagram).

Moving image files comprise of a number of technical features that can influence the quality and long-term preservation or access of the file itself ([Lucker, et al, 2019](#)):

- **Bit depth** or **colour depth** defines the number of colours a pixel may represent. This is expressed in how many bits are used to define the colour, from 1-bit to 48-bit. For example, 24-bit colour depth represents true colour, or what the human eye can see ([Wikipedia, 2021](#)).
- **Colour space** defines how colours are expressed as numbers. Colour spaces are modelled numerically and standardised through what is known as **colour profiles**. RGB, YUV, and YCbCr are examples of common colour profiles ([Blewer, 2020](#)). Colour profiles can be converted from one to another, but such conversions can alter the quality and display of colours ([ASMP, 2015](#)).
- **Compression** algorithms, both lossy and lossless, can be used to reduce the storage size of a moving image file. Lossy compression reduces quality to achieve smaller file sizes, which leads to an irreversible degradation of audio and image quality. Lossless compression reduces size while retaining all original information, thereby preserving the quality of both the audio and video streams.

## 2 Preservation Challenges

Preserving moving images often requires specialized knowledge about formats, carriers, equipment, and software. Broader knowledge on these topics will allow creators and preservation generalists to understand basic good practice for working with moving images. There is an urgency associated with the migration of formats due to carriers becoming obsolete and the rapid decay of materials.

### 2.1 Timeliness of digitization from analogue carriers

Large volumes of moving image recordings on analogue carriers are at risk of loss due to media degradation and obsolescence as hardware necessary to playback and/or digitize is no longer available or becomes prohibitively expensive ([Casey, 2015](#)).

## 2.2 Analogue material assessment

Analogue moving image material requires access to increasingly rare equipment to view. Even with appropriate and often expensive equipment, playback can damage the physical material, which means that transfers to digital formats must be done with care and precise monitoring. Only minimal review of content prior to digitization can be done using physical attributes, documents or metadata prior to digitization.

## 2.3 Volume

Moving image files tend to be very large. This affects multiple aspects of their preservation:

- Increased storage allocations (and the associated long-term costs) may exceed budgets.
- Storage requirements may exceed institutional capacity or the digital repository's processing capabilities.
- Creating preservation and access derivatives further increases storage requirements.
- Fixity checking, automated quality control, metadata extraction, and other processing tasks may take too long to be practical.
- Checking large transfers for fixity when moving files between storage locations or creating backups is time-consuming.

## 3 File Formats

There is no single perfect format for the preservation and future use of moving images and no clear consensus in the digital preservation community. Decisions made on file formats should be dependent on the features and functionality to be preserved and the future use cases to be supported. Note that the tables below do not provide an exhaustive list of formats suitable for preservation and access. The most suitable format for preserving the important features and functionality of a file may be the original format that it was created in. It is recommended that careful research and analysis is carried out before migrating files to a new format.

| File format                                     | Extensions | Brief summary  |
|---|------------|--|
| Digital Picture Exchange (DPX)                  | .dpx       | DPX ( <a href="#">SMPTE standard ST 268-1:2014</a> ) is a raster image file format intended for storing moving image content as a sequence of single .dpx files, typically created from digitizing analogue film carriers at high resolution and bit depth.  |
| Matroska Multimedia Container/ FFV1 Video Codec | .mkv       | Matroska stores CRC32 checksums for every portion of the container, allowing stream-level integrity checking even if the embedded metadata changes ( <a href="#">Matroska, 2005-2020</a> ). FFV1 stores checksums at the frame or frame slice level ( <a href="#">ffmpeg, 2020</a> ). FFV1 within Matroska is a losslessly compressed format standardized through IETF CELLAR-WG ( <a href="#">IETF, 2021</a> ). |
| Material Exchange Format (MXF)                  | .mxf       | MXF is a video container format that often stores JPEG2000 encoded video ( <a href="#">Library of Congress (LC), 2012</a> ). MXF has also been used as the audio and video packaging format for Digital Cinema Packages (DCP). DCP is a more formal folder structure that  |

|   |            |  |
|---|------------|--|
|   |            | contains MXF files alongside other audiovisual components and playback metadata.   |
| Uncompressed V210 in Quicktime or other container | .mov, .avi | Video can be stored as an uncompressed data stream using the V210 Video Picture Encoding. These streams are most often wrapped inside an Apple Quicktime container file (.mov) or sometimes a Microsoft Audio Video Interleave container file (.avi). The typical video content inside has an 8-bit or 10-bit pixel depth with a 4:2:2 chroma subsampling (Blewer, 2020). The resulting file size is very large. |
| Apple ProRes                                      | .mov       | Apple ProRes is a format intended for intermediary video editing and has been used as a standard archival format for both preservation and access (Apple, 2018). This encoding is proprietary and lossy, but popular for use in mainstream audiovisual production and distribution workflows due to ease of access.  |
| H.264/MP4   | .mp4       | H.264 encoding wrapped in an MP4 container produces a file that is relatively small and is capable of web playback. It is a highly ubiquitous access format.   |
| WebM  | .webm      | The WebM container is an open moving image file format primarily used for web playback. It was extended from the Matroska format by Google (WebM Project, 2010-2021).  |

## 4 Metadata Standards

There are a range of potential standards targeting moving image data types. These are suitable for different purposes, including the capture of descriptive and technical metadata.

- EBUCore is an extension of DublinCore, a descriptive metadata standard, developed by the European Broadcasting Union (EBU, 2020).
- EN 15907 is a European Standard that specifies a set of metadata for the description of cinematographic works and provides terminology for use by parties wishing to exchange descriptive metadata. (Filmstandards.org, 2011). It is adopted by many major film archives, including the BFI National Archive and the Center national du cinéma et de l'image animée (CNC), thanks in part to its use within the [FIAF Moving Image Cataloguing Manual](#) (Fairbairn, et al., 2016).
- PBCore is an extension of Dublin Core developed by United States public broadcasting communities and maintained by the WGBH Educational Foundation (PBCore, 2020).
- VRA Core is a descriptive metadata standard developed and maintained by the Library of Congress in partnership with the Visual Resources Association. This metadata standard focuses on general visual artworks (LC, 2018).

## 5 Tips for creators

Whether creating born-digital moving images or digitizing analogue collections, consider consulting an archive for advice and guidance on current good practice for creation.

### 5.1 Born-digital

- Creators of born-digital moving images should be aware of practices that will ensure their content can be preserved. WITNESS, an organization that promotes use of video to protect and defend human rights, has created a guide on how to authenticate, manage, and preserve video ([WITNESS](#), 2020). This guide contains robust recommendations for creating moving image media and metadata to ensure that files can be accessed widely and preserved with the appropriate context.
- Use the most ubiquitous and web-friendly video file format available for access derivatives. Currently, that format is an H.264-encoded video stream within an MP4 container. The popularity of access formats will change over time.
- If possible choose an open format (for example, FFV1 in Matroska) for creation of a primary preservation copy.
- Avoid creating content using applications or web platforms that lock in content or limit use. Test whether you can extract the content from the application or platform and re-use it elsewhere.
- Create a retention policy describing which content should be preserved. For example, decide and document which of the following should be retained: the final production, the associated raw video, special effects work, outtakes, and metadata describing shot lists or editing order.
- Establish how web streaming content will be captured. If export is not possible, screen recording technology can be used to capture content and associated media elements such as sidebar chats.
- If creating video files from live performances, be familiar with best practices when establishing your streaming video strategy. A good resource is Histories of Performance Documentation ([Giannachi](#), 2017), an anthology of research and conversations with experts in the museum space.

### 5.2 Digitized

Migrating from analogue to digital format is essential for videotape and optical disk carriers due to the high risk of degradation of physical material and obsolescence of mechanical hardware required for playback ([Casey](#), 2015). For film carriers, migration is typically required to provide access, and should also be considered for preservation. For most moving image materials, analogue to digital migration happens in 1:1 time.

- Resources are available to help with initial assessment of analogue carriers:
  - The [Texas Commission on the Arts](#) (2004) has a *Videotape Identification and Assessment Guide* with pictures and lists of distinguishing features for most video formats.
  - [meemoo, Flemish Institute for Archives](#) (2020), provides a questionnaire for determining specific formats based on their physical attributes.
  - AVCompass provides an educational resource on giving “direction in organizing, preserving, seeing and appreciating” audiovisual collections, from a physical inventory standpoint ([Bay Area Video Coalition](#), 2015).

- Prioritize videotape and optical disk digitization projects by the fragility of the medium:
  - Highest risk: ½" Open Reel and 2" Open Reel tapes.
  - High risk: U-matic, DVD, CD, DAT, and MiniDV, Video8 and Hi8.
  - Lower risk: Open reel audio and DV are medium risk, and VHS, audiocassette, and Betacam ([Morel](#), 2019).
- When properly stored, film is more stable than magnetic media, but migration from analogue to digital is still necessary for access and should be considered for preservation.
- Small gauge film formats (e.g. 8mm) are at higher risk and should be given digitization priority. However, the physical condition of the film carriers is the greatest factor, and the presence of vinegar syndrome, mould, and similar conditions should lead to assigning affected holdings highest priority for digitization.

## 6 Tips for archivists

### 6.1 General guidance

The following resources provide guidance on preserving and providing access to moving images:

- DPC Technology Watch Reports, *Pragmatic Audiovisual Preservation* ([Blewer](#), 2020) and *Preserving Moving Pictures and Sound* ([Wright](#), 2012).
- For more information on the audio data type, please see the Guidance Note on Audio.
- *IASA-TC 06 Guidelines for the Preservation of Video Recordings* ([IASA](#), 2019).
- Digital Archiving of Film and Video: Principles and Guidance ([Memoriav](#), 2019).
- UW Preservation: Moving Images LibGuide ([University of Washington Libraries](#), 2020).

### 6.2 Community assistance

Engage with the community for advice and support. Professional communities provide guidance and recommendations through published recommendations, working groups, conferences, and online groups.

- The Association of Moving Image Archivists ([AMIA](#)) offers conferences and other events based in the U.S., task forces and working groups, student chapters, and webinars.
- The International Federation of Film Archives ([FIAF](#)) offers training events, a technical symposium, databases, and an academic journal.
- The International Association of Sound and Audiovisual Archives ([IASA](#)) offers a mailing list, conference, and an academic journal.
- The [No Time To Wait](#) events (held annually and always streamed online) provide a forum for information exchange on the digital preservation of audiovisual collections.

### 6.3 Characterization

Characterization can be useful to identify file formats, extract metadata, identify broken or encrypted content, or check conformance to profiles or standards. Tool support and effectiveness can vary considerably for different file formats.

- File format identification can be performed with a tool such as [DROID](#), [FIDO](#), or [Siegfried](#) that uses the [PRONOM file format registry](#). Note that this may only identify the container format. Tools such as [MediaInfo](#) (MediaArea, 2020) can be employed to identify encoded data streams.
- Use MediaConch to validate Matroska containers.

## 6.4 Preservation action

- A discussion of data migration, challenges, and solutions can be found in *On the Bright Side of Data Migrations* (Kromer, 2018).
- Moving images often contain more than one data stream. Careful analysis of each one is essential when undertaking any kind of preservation action, such as making derivative files or performing migration. Consider audio, subtitle tracks, embedded metadata, and other data streams to be as important as the video stream.
- Research into errors and unexpected results can begin with the community-driven research database, *A/V Artifact Atlas* (2020) or the expansive and highly visual reference book, *Compendium of Image Errors in Analogue Video* (Gfeller, et al, 2012).

## 6.5 Quality assurance

- Check data for completeness, including evidence of successful fixity verification and of checks of aural and visual quality. *Carnegie Hall* (2018) outlines quality assurance practices and provides sample scripts for bulk operations. The *UCLA Library Special Collections* (2015) metadata quality assessment workflow can be used as a guide.
- Full quality assurance requires at least as long as the playback duration of the material.
- Contracts with external vendors typically require the vendor to reprocess files with errors during a specified period of time, and quality assurance assessments must be completed during that period.

# 7 References

Apple (2018) *About Apple ProRes*. Available at:

<https://web.archive.org/web/20201205215101/https://support.apple.com/en-us/HT202410>

ASMP (2015) *Best practices: Color Space and Colour Profiles*. Available at:

<https://web.archive.org/web/20201001053019/https://dpbestflow.org/color/color-space-and-color-profiles>

A/V Artifact Atlas (2020) *A/V Artifact Atlas*. Available at:

<https://web.archive.org/web/20200808170723/http://www.avartifactatlas.com/index.html>

Bay Area Video Coalition (2015) *AVCompass*. Available at:

<https://web.archive.org/web/20201001083015/http://www.avcompass.bavc.org/>

Blewer, A. 2020. *Pragmatic Audiovisual Preservation Technology Watch Report: Digital Preservation Coalition*. Available at: <http://doi.org/10.7207/twr20-10>. DOI: 10.7207/twr20-10

Carnegie Hall (2018) *quality-control*. Available at:

<https://web.archive.org/web/20201109054957/https://github.com/CarnegieHall/quality-control>

Casey, M. (2015) 'Why Media Preservation Can't Wait: The Gathering Storm' *IASA Journal* 44, pp. 14–22. Available at: <https://web.archive.org/web/20200906155506/https://mdpi.iu.edu/doc/storm.pdf>

EBU (2020) *Metadata Specifications*. Available at:

<https://web.archive.org/web/20201113142206/https://tech.ebu.ch/MetadataEbuCore>

- Fairbairn, N., Pimpinelli, M.A., & Ross, T. (2016) *The FIAF Moving Image Cataloguing Manual*. Available at:  
<https://web.archive.org/web/20200930164358/https://www.fiafnet.org/images/tinyUpload/E-Resources/Commission-And-PIP-Resources/CDC-resources/20160920%20Fiaf%20Manual-WEB.pdf>
- Filmstandards.org (2011) *EN 15907*. Available at:  
[https://web.archive.org/web/20201113080132/http://filmstandards.org/fsc/index.php/EN\\_15907](https://web.archive.org/web/20201113080132/http://filmstandards.org/fsc/index.php/EN_15907)
- FFmpeg (2020) *ffmpeg*. Available at  
<https://web.archive.org/web/20210101002538/https://ffmpeg.org/ffmpeg.html>
- Gfeller, J., Jarczyk, A., and Phillips, J. (2012) *Compendium of Image Errors in Analogue Video*. SIK ISEA, Schweizerisches Institut für Kunstwissenschaft, ISBN: 3858813818, 9783858813817
- Giannachi, G. and Westerman, J. (2017) *Histories of Performance Documentation*. Available at:  
<https://doi.org/10.4324/9781315645384>. DOI: 10.4324/9781315645384
- Harvey, P. (2020) *ExifTool by Phil Harvey*. Available at:  
<https://web.archive.org/web/20201112030540/https://exiftool.org/>
- Kromer, R. (2018) “*On the Bright Side of Data Migrations*”. International Association of Sound and Audiovisual Archives (IASA) Journal, no. 49 (November), 18–22. <https://doi.org/10.35320/ij.v0i49.72>. DOI: 10.35320/ij.v0i49.72
- IASA (2019) *IASA-TC 06 Guidelines for the Preservation of Video Recordings*. Available at:  
<https://web.archive.org/web/20201201202421/https://www.iasa-web.org/tc06/guidelines-preservation-video-recordings>
- IEEE Xplore (2017) *ST 268-1:2014 - SMPTE Standard - File Format for Digital Moving-Picture Exchange (DPX)*. Available at:  
[https://web.archive.org/web/20200815024836if\\_/https://ieeexplore.ieee.org/document/7936730](https://web.archive.org/web/20200815024836if_/https://ieeexplore.ieee.org/document/7936730)
- IETF (2021) *Codec Encoding for LossLess Archiving and Realtime transmission (cellar)*. Available at:  
<https://web.archive.org/web/20210112130234/https://datatracker.ietf.org/wg/cellar/about/>
- Library of Congress (2020-2021) *Recommended Format Statement: III. Moving Image Works*. Available at:  
<https://web.archive.org/web/20201121131208/https://www.loc.gov/preservation/resources/rfs/moving.html>
- Library of Congress (2018) *VRA CORE - a data standard for the description of images and works of art and culture*. Available at:  
<https://web.archive.org/web/20201110173949/https://www.loc.gov/standards/vracore/>
- Library of Congress (2012) *Sustainability of Digital Formats: Planning for Library of Congress Collections: Material Exchange Format (MXF)*. Available at:  
<https://www.loc.gov/preservation/digital/formats/fdd/fdd000013.shtml>
- Lucker, P., Sijtsma, C., and van Veenendaal, R. (2018) *Significant Significant Properties*. Available at:  
<https://osf.io/mc2ea/>
- Matroska (20025-2020) *Matroska Media Container — Homepage*. Available at:  
<https://web.archive.org/web/20210215130908/https://www.matroska.org/index.html>

MediaArea (2020) *MediaInfo*. Available at  
<https://web.archive.org/web/20210120180559/http://mediaarea.net/en/MediaInfo>

meemoo, Flemish Institute for Archives (2020) *What image and sound do you have stored in your (personal) archive?* Available at:  
<https://web.archive.org/web/20201123192918/https://knowyourcarrier.com/>

Memoriav (2019) *Digital Archiving of Film and Video: Principles and Guidance*. Available at:  
<https://web.archive.org/web/20201104121338/https://memoriav.ch/dafv/?lang=en>

Morel, M. (2019) *Bay Area Video Coalition Guide to Audiovisual Preservation*. Available at:  
[https://web.archive.org/web/20200930220934/https://bavc.org/sites/default/files/wysiwyg-uploaded/BAVC\\_Guide\\_To\\_Audiovisual\\_Preservation\\_2019\\_0.pdf](https://web.archive.org/web/20200930220934/https://bavc.org/sites/default/files/wysiwyg-uploaded/BAVC_Guide_To_Audiovisual_Preservation_2019_0.pdf)

PBCore (2020) *PBCore Metadata Standard*. Available at:  
<https://web.archive.org/web/20200216053700/http://pbcore.org/>

Society of American Archivists [SAA] (2020) *Dictionary of Archives Terminology*. Available at:  
<https://web.archive.org/web/20200807182019/https://dictionary.archivists.org/entry/moving-image.html>

Texas Commission on the Arts (2004) *Videotape Identification and Assessment Guide*. Available at  
<https://web.archive.org/web/20210208100747/https://www.arts.texas.gov/wp-content/uploads/2012/04/video.pdf>

UCLA Library Special Collections (2015) *Metadata Quality Control Guidelines*. Available at:  
[https://web.archive.org/web/20150418081458/https://www.library.ucla.edu/sites/default/files/Guidelines\\_MetadataQualityControl.pdf](https://web.archive.org/web/20150418081458/https://www.library.ucla.edu/sites/default/files/Guidelines_MetadataQualityControl.pdf)

University of Washington Libraries (2020) *Preservation: Moving Image LibGuide*. Available at:  
<https://web.archive.org/web/20201204211237/https://guides.lib.uw.edu/c.php?g=342245&p=2305091>

WebM Project (2010-2021) *The WebM Project | Welcome to the WebM Project*. Available at:  
<https://web.archive.org/web/20210101010444/https://www.webmproject.org/>

Wikipedia (2021) *Color Depth*. Available at:  
[https://web.archive.org/web/20201203060637/https://en.wikipedia.org/wiki/Color\\_depth](https://web.archive.org/web/20201203060637/https://en.wikipedia.org/wiki/Color_depth)

WITNESS (2020) *WITNESS | Human Rights Video*. Available at:  
<https://web.archive.org/web/20200903143046/https://www.witness.org/>

Wright, R. (2012) *Preserving Moving Pictures and Sound*. Available at:  
<http://dx.doi.org/10.7207/twr12-01>. DOI: 10.7207/twr12-01