

# Managing digital records: 5.5

---

## Contents

5.5	Managing digital audio and video recordings .....	1
	Digitisation of analogue audio and video.....	1
	The difference between analogue and digital audio and video .....	1
	Advantages of digital over analogue .....	2
	Why digitisation of analogue audio and video recordings may be required .....	2
	Prioritising digitisation of analogue audio and video recordings .....	3
	Managing a digitisation project .....	3
	Digital audio and video file formats.....	4
	Metadata for the digitisation of analogue recordings.....	4
	‘Essential characteristics’ of analogue recordings .....	7
	Disposal of analogue source records .....	7
	Preservation of digital files .....	7
	Quick tips for digitising analogue formats .....	8
	Case study: Digitisation on a grand scale - The National Film and Sound Archive.....	8
	Case study: Preserving oral history - Department of Services Technology and Administration’s digitisation project .....	9
	Outsourcing the digitisation of audio or video.....	10
	Digital audio and video file formats.....	12
	Choosing a format/codec for creating digital audio or video.....	13
	Choosing a format/codec for digitising analogue audio or video .....	15
	Criteria for choosing formats for preservation purposes .....	17
	Formats and codecs for digital audio preservation.....	18
	Formats and codecs and formats for digital video preservation .....	22
	Case study: Tourism NSW - Use of digital video to promote the State .....	24

## 5.5 Managing digital audio and video recordings

There are number of issues to consider regarding managing and preserving audio-visual records over time. We have developed guidance on the digitisation of analogue audio and video and on what digital audio and video file formats are suitable for the creation and preservation of audio-visual records.

### Digitisation of analogue audio and video

#### The difference between analogue and digital audio and video

Analogue recording is a linear process involving the creation of variations in a recording medium that correspond to variations in the signal being captured. Recording media used for this purpose include record albums and audio and video cassette or reel-to-reel tapes.

Analogue devices, such as VCRs, tape and record players, read analogue media by physically scanning these variations. For example, a record player creates an audio signal by translating the bumps and dips in the grooves of an album and a tape player creates an audio or video signal by reading the variations in the intensity of magnetisation of the tape.

Digital recordings are made using the binary system. Digital devices such as CD recorders and camcorders convert the signal and turn it into digital information – a sequence of numbers – sampling it at set intervals. The higher the sampling

rate, the more accurate the conversion is, which translates into higher quality sound or video.<sup>1</sup>

### Advantages of digital over analogue

Both analogue and digital methods have advantages and disadvantages. See Wikipedia at [http://en.wikipedia.org/wiki/Analog\\_recording\\_vs.\\_digital\\_recording](http://en.wikipedia.org/wiki/Analog_recording_vs._digital_recording) regarding analogue recording versus digital recording for more information.

The main advantages of digital over analogue methods of recording and playback include:

- the binary code of digital audio and video can be read by a computer. Therefore, computers can be used to edit the data e.g. remastered to enhance sound and visual quality, and to create new effects. Analogue signals can only play what was originally recorded as it was recorded
- digital media is non-linear (or non-real time), so it can be edited and played back starting at any point
- digital information does not degrade and lose quality with repeated use (like tapes or record albums do). They may be copied repeatedly without loss if they are not re-encoded
- groups of numbers can be compressed by finding patterns in them meaning the same information can be stored more efficiently.<sup>2</sup>

### Why digitisation of analogue audio and video recordings may be required

Digitisation refers to the conversion of non-digital material to digital form. Reasons why your organisation may consider converting existing analogue recordings to digital include:

- you face difficulties in providing access to existing resources as the devices to read them are no longer widely available in the organisation or elsewhere (for example, record players or VCRs)
- your existing analogue resources are becoming fragile or being degraded by poor storage conditions, the passage of time or overuse, threatening their accessibility
- you wish to improve the recordings e.g. digitally enhance them or improve indexing of the recordings and therefore search and retrieval (note: an enhanced version should not replace the original unenhanced master but can be an alternative for delivery)
- you wish to make them more readily available by providing the potential for online delivery
- you wish to create new versions of the recordings e.g. tailor them for other uses.<sup>3</sup>

For more information about the factors to consider before embarking on a digitisation project, see the JISC fact sheet, *Deciding to digitise* at <http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/deciding-to-digitise/>.

---

<sup>1</sup> Sharpened.net Help Center: *What is the difference between analog and digital technology?* Available at: <http://www.sharpened.net/helpcenter/answer.php?62>

<sup>2</sup> How stuff works: *Can you explain the basic difference between analog and digital technology?* Available at: <http://electronics.howstuffworks.com/question7.htm>

<sup>3</sup> JISC Digital Media, *Deciding to digitise*, 14 November 2008, available at: <http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/deciding-to-digitise/>

## Prioritising digitisation of analogue audio and video recordings

All public offices have an obligation to retain accessibility to technology/equipment dependent records under s.14 of the State Records Act. This applies to all State records, regardless of how long they are required to be kept. Digitisation may sometimes be required to guarantee continued accessibility.

However, digitisation of analogue audio and video can be very resource intensive, depending on the quantities and formats involved, the equipment and software required and the quality of the original analogue source. Your organisation may, by necessity, need to take a risk based approach and establish priorities for digitisation.

Priorities should include analogue recordings that:

- are masters and are used very frequently (as analogue recordings become degraded by use. Generally copies should not be made from masters)
- are already being degraded by high use and therefore in danger
- have the disposal action as 'retain in agency' or 'required as State archives' in the organisation's current retention and disposal authority (as these indicate the recordings that have significant ongoing interest to the organisation or the community).

State Records has produced a *General retention and disposal authority – audio visual programs and recordings* (see <http://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/rules/general-retention-and-disposal-authorities/audio-visual-programs-and-recordings-gda-11/audio-visual-programs-and-recordings-gda-11>) for audio, film and video programs and recordings created or commissioned by NSW public offices from c.1966 onwards. This will indicate what audio-visual records are required as State archives. The organisation's functional retention and disposal authority may also contain additional information about disposal requirements for audiovisual materials created by the organisation.

The JISC fact sheet *Selection procedures* may assist you to make further selection decisions. See <http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/selection-procedures-for-digitisation/>

## Managing a digitisation project

If your organisation needs to convert analogue audio or video to digital, you will need to consider very carefully whether to outsource the digitisation or to manage it internally.

Running an internal project has certain advantages including that:

- the organisation acquires the equipment and develops staff expertise making it available for future digitisation projects
- the movement and treatment of materials can be closely supervised
- the procurement and tendering process can be avoided, which is often expensive in its own right
- staff time and some other overheads can be borne by the organisation rather than being an additional visible cost.

However, the advantages of using a contractor may be:

- the organisation avoids the cost of obtaining expensive equipment and procuring expertise or training staff

- the work can be done according to strict deadlines and it will not impact as greatly on staff time
- the contractor can provide the necessary space.

Sometimes it is possible or preferable to outsource only part of the project. For example, the British Library outsourced the digitisation element of an audio project but managed the selection of content, digital rights management and metadata creation issues internally.<sup>4</sup>

Advice to assist you in making this decision is provided in the JISC factsheets, *To outsource or to digitise in-house?* at <http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/to-outsource-or-to-digitise-in-house/>

Appendix 2 contains a list of issues to consider if you are considering outsourcing the digitisation of audio or video.

If you are considering internally managing a project, the JISC *Digital media website* also has an extensive range of resources on digitisation of analogue audio and video. For example, they address budgeting for a project, project management and quality assurance. For a full list see the JISC information on *cross media*. [Link to <http://www.jiscdigitalmedia.ac.uk/crossmedia/>]

### Digital audio and video file formats

The purpose of the project, requirements for video and audio quality and the length of time files need to be retained (as specified in an authorised retention and disposal authority) need to be considered when selecting an appropriate file format. For more information see Digital audio and video file formats.

### Metadata for the digitisation of analogue recordings

Recordkeeping metadata is structured information that describes a record. It includes information about the record's structure, the context in which it was created and its content. Metadata is valuable in searching for, retrieving, accessing and using the record, but it also helps to verify its authenticity and promote its long term management and preservation.

The *Standard on digital recordkeeping* (see <http://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/rules/standards/standard-on-digital-recordkeeping>) outlines the minimum requirements for recordkeeping metadata. Your digitised recordings must be described by this metadata. Your organisation is likely to derive some of this information from existing metadata kept about your analogue files. For example, you might have information on a record album cover describing the recording e.g. place of recording, event, speaker etc or a reel-to-reel video might have a label showing the title, production details, date made and viewing restrictions. Additional information may be stored in the organisation's databases or in separate data files.

As part of preliminary stages of a digitisation project your organisation should identify any additional requirements for metadata (beyond the minimum required) and draw up clear specifications. Examining existing metadata schemas

---

<sup>4</sup> S Tanner, *Cost Reduction in Digitisation*, Minerva Plus Project, Version 1, June 2006, available at: [http://www.minervaeurope.org/publications/CostReductioninDigitisation\\_v1\\_0606.pdf](http://www.minervaeurope.org/publications/CostReductioninDigitisation_v1_0606.pdf)

for audio visual materials (e.g. METS or SMIL)<sup>5</sup> or those used for similar industries/projects may be of benefit.

The following table describes some metadata that may be useful to capture (where relevant and available). The second column indicates whether this is part of the minimum required metadata specified in the *Standard on digital recordkeeping*.

<b>Identifiers</b>	<b>Minimum required metadata in <i>Standard on digital recordkeeping</i>?</b>
Title e.g. program, segment, episode or interview title; e.g. title of album, title of tracks on album	Yes. A repeated title field and an encoding scheme may be required to specify the type of title.
Date e.g. date of recording, broadcast, production	The date of creation (recording) is minimum required metadata. A repeated date field and an encoding scheme may be required to specify the type of date.
Unique identifier e.g. item number, Barcode, unique material identifier	Yes. A repeated unique identifier field and an encoding scheme may be required to specify the type of unique identifier.
<b>Creator information</b>	Who/what created the record is minimum required metadata.
Credits e.g key individuals who made the work. In commercial productions could include producer, director, production company, producing agency	Who/what created the record is minimum required metadata. An encoding scheme may be required to specify the type of individuals/organisations involved.
Personnel e.g. presenter, interviewer, interviewee	Who/what created the record is minimum required metadata. An encoding scheme may be required to specify the type of personnel involved.
<b>Content information</b>	
Genre or style of recording e.g. interview, report, live-to-air, field recording, oral history	Yes. The record type is required.
Content note e.g. topics in recording, people in recording, summary	Not required metadata in the Standard but may be useful to capture for business purposes and to promote access.
<b>Physical or technical characteristics</b>	
Quantity e.g. audio cassette 2 of 2	Not required metadata in the Standard but may be useful to capture for business purposes and to promote access..
Duration e.g. feet of moving images or time	Not required metadata in the Standard but may be useful to capture for business purposes and to promote access..
Class of audiovisual material e.g. film	Yes. The creating application is

<sup>5</sup> For more information, see the metadata section of 'Audio and Video Capture and Management' in the National Initiative for a Networked Cultural Heritage publication, *The NINCH guide to good practice in digital representation and management of cultural heritage materials* Version 1. October 2002, available at: <http://www.nyu.edu/its/humanities/ninchguide/VII/>

reel, magnetic tape, optical disc	minimum required metadata. This would be made up of at least the medium and the file format if applicable.
Specific type of component e.g. BWF file on Hi 8 tape, original negative, digital beta	Yes. The creating application is minimum required metadata. This would be made up of at least the medium and the file format if applicable.
Physical dimensions of carrier e.g. 12" Vinyl, 7" magnetic reel-to-reel tape	Not required metadata in the Standard but may be useful to capture for business purposes and to promote access..
Gauge (for motion picture film – e.g. 8mm, 6mm)	Not required metadata in the Standard but may be useful to capture for business purposes and to promote access..
Colour characteristics e.g. black and white, colour	Not required metadata in the Standard but may be useful to capture for business purposes and to promote access..
Sound characteristics e.g. silent, mute, playback speed, number of tracks (e.g. 4 track audio cassette recording), encoding (Dolby, stereo, mono) <sup>6</sup>	Not required metadata in the Standard but may be useful to capture for business purposes and to promote access..

Further metadata will be required when you convert the analogue recording to a digital format. For example, metadata will need to be captured about the digitisation process itself such as the date of digitisation and who/what undertook it (required by the *Standard on digital recordkeeping*), and about the old and new formats. The National Film and Sound Archive (NFSA) recommend the collection of the following minimum metadata for preservation masters:

- Digital file recorder configuration settings
- Digital file recorder identification; type, version, serial number
- Details of the source tape and identifier
- Source videotape reproducer identification; type, serial number etc
- Master preservation file encoding details; codec type, versions etc
- Transfer operator identification
- Video and audio signal analysis metrics.<sup>7</sup>

Some metadata, particularly some technical information, can be automatically captured and embedded in some digital file formats or added later using an editor. Wherever possible your organisation should verify that any automatic metadata capture is correct and records full technical information, and not just an edited or genericised format name.

As digital audio and video are subject to rapid change, and technological obsolescence is a real threat, a review date should also be included in the metadata so that the organisation is reminded about the need to monitor its condition and useability and if refreshment or replication is needed.

<sup>6</sup> R Gamble and L. Curham, Chapter 17: 'Sound recordings' in *Keeping Archives*, Australian Society of Archivists, Third edition, 2008 p.555-556.

<sup>7</sup> National Film and Sound Archive (NFSA), *Digital media preservation: Video* [unpublished]

If after digitisation the recordings (and/or the analogue source records) are to be sent to State archives, the organisation should also ensure that metadata such as box listings and references to supporting documentation e.g. transcripts, release forms are adequate and in line with State Records' transfer requirements. State Records should be contacted prior to transfer.

For more information about metadata, refer to JSIC fact sheet *An introduction to metadata* at <http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/an-introduction-to-metadata/> which also contains links to more extensive information. See also State Records NSW, *Managing digital records, 3. Use recordkeeping metadata for digital recordkeeping*.

### **'Essential characteristics' of analogue recordings**

All records have what is known as 'essential characteristics' (also referred to as 'significant properties' or 'essence'). For digital records, essential characteristics are those properties that must be preserved over time and domains and across changing technological environments to ensure the continued authenticity, accessibility, useability and meaning.<sup>8</sup>

When digitising analogue audio and video or migrating digital information it is important to ensure that their essential characteristics are identified e.g. with a digital video is colour, interactivity or sound important? With digital audio, can you hear the recording in its entirety? Are breaks correctly inserted so that tracks do not overlap or run together? Testing/quality control built into the digitisation process should confirm that when digitised these essential characteristics are faithfully reproduced. With digital video, in particular, studies have shown that the essential characteristics can be reproduced in digitisation providing the right file format is chosen.<sup>9</sup>

### **Disposal of analogue source records**

Source records remaining after digitisation of analogue materials have been excluded from the *General retention and disposal authority – Source records that have been migrated* (see <http://www.records.nsw.gov.au/recordkeeping/government-recordkeeping-manual/rules/general-retention-and-disposal-authorities/source-records-that-have-been-migrated-ga-33/source-records-that-have-been-migrated-ga-33>). As yet State Records has not authorised a retention and disposal authority to cover source records from analogue to digital formats. Contact State Records if you need assistance regarding the disposal of source records.

### **Preservation of digital files**

The preservation of the digital files does not end with their creation and the capture of metadata. Digital audio and video formats have their own preservation issues. For example, rapid obsolescence of digital technology and media instability makes digital media particularly vulnerable to loss (most have a life span of 5-10 years). Your organisation will need to monitor the files regularly to ensure they remain accessible and useable. Preservation may involve regular refreshment or regular replication to formats that can maintain their 'essential

---

<sup>8</sup> G.Knight, 'Same as it ever was: significant properties and the preservation of meaning through time' presented at *Decoding the Digital: A common language for preservation*, 27 July 2010, available at: <http://www.dpconline.org/events/decoding-the-digital-a-common-language-for-preservation.html>

<sup>9</sup>M Coyne and M Stapleton, *The significant properties of moving images*, 26 March 2008, available at: [http://www.jisc.ac.uk/media/documents/programmes/preservation/spmovimages\\_report.pdf](http://www.jisc.ac.uk/media/documents/programmes/preservation/spmovimages_report.pdf)

characteristics'. The process of shifting from one digital format to another is often referred to as 'transcoding'. See the JISC fact sheets *Transcoding digital video* at <http://www.jiscdigitalmedia.ac.uk/movingimages/advice/transcoding-digital-video/> and *An introduction to digital preservation* at <http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/an-introduction-to-digital-preservation/>

The International Association of Sound and Audiovisual Archives (IASA) have also developed principles regarding the digital preservation of audiovisual materials. See the IASA website at [www.iasa-web.org](http://www.iasa-web.org).

Note: For a full listing of the JISC Digital Media fact sheets available regarding digitisation projects see <http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/overview-of-materials-for-digitisation-project-staff/>

### Quick tips for digitising analogue formats

- Treat the best version of the analogue recording as an analogue master.
- Digitise in the highest quality format you can. Use no compression (or lossless compression if compression is unavoidable) for creating a digital preservation master. Lossy compression irretrievably reduces quality. Use open formats or formats with good 'sustainability factors' where possible.
- Define your metadata requirements carefully and build metadata collection into the project.
- Build testing into the digitisation project to ensure the 'essential characteristics' are maintained.
- If enhancements are to be made, make these from a copy. Save a 'master' of the original and enhanced version.
- Generate lower resolution copies for delivery (generally these will be produced from the enhanced version).
- If possible, have another copy of the analogue master, preservation master and enhanced master stored at a different location for disaster management purposes.
- Regularly monitor your digital masters for obsolescence or degradation.<sup>10</sup>

### Case study: Digitisation on a grand scale - The National Film and Sound Archive

The National Film and Sound Archive recently announced that it has purchased a SAMMA digitisation system to use JPEG 2000 frame images, losslessly compressed with an MXF wrapper for the digitisation of more than 40,000 hours of video.

JPEG2000 is a standard open encoding format that has growing support within the audiovisual and archival communities. It enables LPCM encoding which is a universal standard for audio encoding. The JPEG2000 lossless compression is completely reversible with no loss to data quality and offers significant data storage savings compared to uncompressed linear encodings.

The Material eXchange Format (MXF) conforms to open, published standards. It works with a variety of compression strategies. It is driven by user needs and has

---

<sup>10</sup> JISC Digital Media, *Generic video digitisation workflow*, 27 Jan 2009, available at: <http://www.jiscdigitalmedia.ac.uk/movingimages/advice/generic-video-digitisation-workflow/>

a strong commercial base. It is widely adopted for content exchange and archival applications. A metadata wrapper describes the material contained within the MXF file.

JPEG2000, MXF and linear pulse code modulation (LPCM) audio are interoperable with a number of digital preservation and production tools available.<sup>11</sup>

### **Case study: Preserving oral history - Department of Services Technology and Administration's digitisation project**

The Department of Services, Technology and Administration implemented a digitisation project to convert around 550 audio tapes to digital format.

#### **Background**

In the lead up to Australia's Bicentennial there was a willingness to invest in projects to capture the nation's history. Some key managers in the Department's predecessor agencies rightly identified the wealth of historical knowledge within their own organisations and established an oral history program to capture it.

The program was initially involved in capturing oral histories relating to the architectural and engineering history of the then Department of Public Works. It was soon expanded and became thematically based, dealing with such diverse themes as the Newcastle Earthquake, the Interior Design Unit, Flood Mitigation in the Hunter Valley, the building of schools and the stonework program to restore sandstone buildings. A professional historian was employed to conduct interviews and summarise themes on a separate audio tape. By 2003/4 between 300 and 400 tapes were in the collection including 8 summary tapes.

The Department also took in oral history collections of some other organisations if they had some involvement with the work undertaken. For example, the collection contains Roads and Traffic Authority oral history interviews concerning the construction of the Sydney Harbour Bridge (a Public Works project). It also contains copies of interviews organised by the Department of Housing in 2005/6 with residents of the historic houses in Millers Point (as the houses were being restored by the Department). The masters of the Millers Point tapes are held by the State Library of NSW. To date there are around 550 tapes.

#### **Decision to digitise**

A Departmental representative contacted State Records to organise the transfer of the audio tapes to the State archives. It was recommended that they first be copied into a stable long term digital format for ongoing preservation and access, as audio tapes and players are becoming obsolete. The format chosen was FLAC (Free Lossless Audio Codec), selected for its lossless and open characteristics, suitable for archival value records.

The Department investigated the costs of digitisation with an external provider and compared this to cost of conducting an in-house digitisation project. The latter was deemed to be more cost effective considering the number of tapes so they purchased and installed the equipment and organised training and procedures.

At the end of the project State Records will receive the digital masters of the recordings and the masters of the analogue tapes along with copies of the documentation e.g. signed releases from interviewees. A backup copy will be retained by the Department for reference.

---

<sup>11</sup> NFSA, Op.cit. See also the *Image and Data Manager* article 'NFSA acts to preserve video', 13 July 2010, available at: <http://idm.net.au/article/007951-nfsa-acts-preserve-video>

## Digitisation project

The project started in January 2010. The Records Project Team within Information Services planned the process carefully and established rigorous procedures, including testing, monitoring and quality control procedures. Documentation created by the team included: an Issues Register established to record any technical issues faced and how they would be resolved; a register to record metadata about the analogue tapes, the digitisation process and who was involved with digitisation; and, daily reports regarding cleaning and running of the tapes and equipment. Strict naming conventions and file locations were also established for the digital files.

At the time of writing around 450 of the tapes have been copied on both sides, with 100 left to digitise.

### Lessons learnt

The team have reported that the project has been a very positive experience. They worked well together and carefully planned all stages of the project including the documentation required. Through regular project meetings they have efficiently identified and managed any problems arising. Significant cost savings have been made by managing the project internally rather than outsourcing.

There were a few issues with the technology. The initial integration of the equipment with the DSTA network caused concerns and network speed was another problem. However, these were soon overcome and few difficulties have been recorded since. A good relationship with IT has proven to be vital to the success of the project.

There have been some challenges associated with the FLAC format as it is not widely supported. For example, it does not work using some standard desktop applications. However, the software used for the project can easily convert the digital file in FLAC format to other formats which have wider compatibility for distribution, web publishing etc.

Material is now coming 'out of the woodwork' so while this particular project will end shortly, other digitisation work may continue in the Department.

### Other digitisation projects

See also JISC *Learning lessons from other digitisation projects*, at <http://www.jiscdigitalmedia.ac.uk/crossmedia/advice/learning-lessons-from-other-digitisation-projects/>

## Outsourcing the digitisation of audio or video

### Before contacting potential contractors

Before contractors are even contacted it is a good idea to identify the following. These issues might be defined in a business case for digitisation/outsourcing:

- the drivers and benefits of digitisation
- the risks identified
- the scope, volume and type of material selected for digitisation
- the condition of the material for digitisation and whether conservation or cleaning is required prior to digitisation
- the digital formats and quality required
- stakeholder interests
- the metadata and documentation required

- ownership, copyright, privacy and digital rights management issues
- what quality control and assurance parameters need to be in place
- insurance requirements for rare or valuable materials
- financial and human resource issues, including roles/skills of staff in monitoring the digitisation process.

### **Before choosing the contractor**

Contractors should be assessed carefully. Some questions to ask may include:

- Can you manage the range of analogue formats we need to digitise?
- Are you able to provide the required digital formats (including suitable archival formats when required)?
- Do you have experience in dealing with these materials/formats (and do you have experience in dealing with unique materials)?
- What are your transportation and materials handling procedures? Is an alternative (for rare, valuable material) to digitise on-site?
- Are you able to provide the required metadata and documentation?
- What quality control, treating and assurance measures do you provide? Do you agree to our proposed monitoring process?
- Are we able to inspect samples of your past work and contact some of your clients? (these should be inspected to determine if of sufficient quality)
- What are the costs and timeframes involved? Are we protected from paying more if costs exceed expectation? What recourse is there if the timeframes are not met?
- What backup and disaster recovery arrangements do you have?

### **Inclusions in the contract**

Contracts should clearly articulate:

- roles and responsibilities
- costs, timeframes and expectations: including that if costs are exceeded this is borne by the contractor and if recourse if timeframes are exceeded
- transportation and materials handling requirements (to prevent damage to the originals)
- the range, type and quantity of material to be digitised
- the digital formats required, including formats of preservation masters (unenhanced) and any enhanced copies or access copies required
- specifications regarding the use of compression (which, if used, should be lossless for archival records)
- additional specifications e.g. sampling, bit rate
- the testing and quality control and assurance procedures to be used by the contractor
- the agreed process of monitoring and quality control and assurance by the organisation
- the metadata and documentation requirements
- where ownership of all materials lies (e.g that ownership of originals and all copies resides with the State)

- arrangements for the return of the analogue and digital recordings to the organisation
- backup and disaster recovery arrangements.

## Digital audio and video file formats

### Definitions

The terminology used for digital audio and video can be extremely confusing, particularly as terms are often used interchangeably. The major terms are defined below.<sup>12</sup>

### Coding and encoding:

The way binary numbers (1 and 0) are used to represent sound or images: the binary numbers can be 'read' by a computer. All video and audio on the internet and in computing systems are encoded in some way.

'Encoding' is often used to refer to both the format of video/audio streams and the file type of digital video or audio files. For example .mp3 and .dv are often referred to as 'encodings' but the former is an audio file format and the latter an encoding format for digital video streams. Some encoding formats are 'wrapper' formats (see Wrapper format below).

### File format:

The structure of a discrete chunk of digital data is known as a file. A file format is a way of capturing the encoded data stream as a logical unit and storing it. Codec software is then needed to interpret the file (see Codec below). Some file formats are known as wrappers.

### Wrapper (or packaging) format:

A wrapper is a file format for encapsulating different bitstreams (or files) and metadata into a single file. For example, in an audio 'wrapper format' like WAVE (.wav), audio streams or files and other chunks of metadata are encapsulated into a single file. AVI, the standard Windows Media wrapper, and MOV, the standard QuickTime container, encapsulate audio, video, metadata, and other control information together as a single package. MXF is also a wrapper format that wraps video, audio and other bitstreams.

### Codec:

Short hand for (en)coder/decoder or compression/decompression. A codec is the method by which an audio or video bitstream is encoded and decoded. For some uses it is important to reduce the size of the data streams e.g. for Internet delivery, so many codecs compress audio and video data. Most of the publicly documented audio file formats can be created with one or more codecs. Without the correct codec software to decode the encoded bitstream, an audio or video file may not play.

Codecs are usually:

- Lossy: A lossy codec discards certain portions of the signal to achieve a smaller file size. For example, mp3 codecs attempt to identify and remove portions of the signal that would not result in a perceived loss of quality of the sound to make the file smaller. The more compression, the more data

---

<sup>12</sup> All definitions are sourced from Arts and Humanities Data Service (AHDS), *Digital moving images and sound archiving study*, August 2006, available at: <http://www.ahds.ac.uk/about/projects/archiving-studies/moving-images-sound-archiving-final.pdf>

is removed and more reduction in audio quality. Once removed, the data cannot be recovered.

- Lossless: A lossless codec achieves smaller file sizes through other methods rather than removing data. For example, a variable bit rate may use fewer bits to encode silences as compared to an active section of music. Lossless codecs do not compromise the audio or video quality even when they hold compressed data.<sup>13</sup> A lossless compressed format requires much more processing time than an uncompressed format but is more efficient in space usage.

Some codecs may be configured to be either lossy or lossless.

### **Choosing a format/codecs for creating digital audio or video**

File formats for creating new audio and video material are often closely associated with the type of equipment used. For example, most cameras will only support a few digital video file formats.

If your organisation has the opportunity to purchase new equipment for recording, they should examine what file formats or codecs are supported and if they are suitable for the intended use and longevity of the recording.

Consider:

- Is the file being created/received in the course of official business and might it be needed to be accessed for use or to provide evidence of business? ie is it a corporate record?
- Is high audio or video quality required for this project? (the quality achieved will be dependent on the equipment used to record as well as the format chosen)
- Is the file being created expressly to meet a short term need? Or does it have significant business value (e.g. intellectual or financial)? Might it have long term cultural or other value? Your organisation's retention and disposal authority will assist you in determining the value of the file and how long it may need to be retained.

### **Short term**

Even if the file is a corporate record, if it only needs to be kept for a very short period of time your organisation should use the format and codec which is most suitable for the aims of the project and compatible with editing software to be used. The processing power and storage space you have available may also impact on your decision.

Example: Many of the MPEG file formats (e.g. mp3) which are standard definition, affordable and widely available may be used if the recordings are only required in the short term and fast delivery (e.g. over the web) is the goal. However, they use lossy compression and are not considered to be of sufficient quality for preservation.

### **Long term or archival, or where quality is paramount**

If the audio or video file:

- will require long term or archival retention or
- needs to be of optimal quality

---

<sup>13</sup> National Archives and Records Administration (NARA), *Frequently asked questions about digital audio and video records*, available at: <http://www.archives.gov/records-mgmt/initiatives/dav-faq.html>

your organisation should try to use a suitable format of as high a quality as possible for the creating of the preservation 'master' file. Duplicating copies and reference copies can then be made at a quality fit for purpose.

Formats should meet as many of the identified Criteria for choosing formats for preservation purposes as possible. However, sometimes these may need to be balanced with the resources and equipment available.

Only linear uncompressed codecs or mathematically reversible lossless codecs should be used for preservation.<sup>14</sup> No compression is ideal, but the files may take up considerable storage space (particularly video files). Your organisation will also need the processing power to be able to open and view the files and do error checking etc.

If your organisation needs to use existing equipment and is therefore tied to the file formats associated with these, they should consider creating a preservation 'master' by copying to a suitable format. Copying to a higher quality format will not increase the quality of the original recording. However, the original recording should never be copied to a format of lesser quality when the aim is to create a preservation master (e.g. do not copy a Digital Betacam produced professionally onto a DVD to create a preservation master). The aim should be to preserve as much quality of content as possible from the original recording.

### Audio

Widely recognised and accepted formats for preservation purposes for audio are WAV or BWF. These are uncompressed and generally use the lossless Linear Pulse Code Modulation (LPCM) codec. They retain as much information as possible from the recording source. Most digital audio capture software can record into these formats. BWF has the advantage over its predecessor WAV of significant metadata capacity and has widely adopted standards for metadata exchange and interoperability.

FLAC is a suitable format for digital preservation as it is stable, free and open source and employs lossless compression. The National Archives of Australia converts a number of other audio files (including WAV) to FLAC during the process of normalisation for storage in their digital archive.<sup>15</sup> If considering this format for business use, however, it is important to note that it is not widely supported by the audio industry.<sup>16</sup>

If the recording is to be optimised, lossless methods should be used. The optimised audio file and information regarding the optimisation should be saved, separately from the preservation master.

For more information and issues to consider when choosing an audio file format, including recommendations in relation to different methods of delivery, refer to the JISC fact sheet Choosing a digital file format at <http://www.jiscdigitalmedia.ac.uk/audio/advice/choosing-a-digital-audio-file-format/>

See also Formats and codecs for digital audio preservation for more discussion about potential formats.

### Video

There is no consensus to date among the archival community as to which file format or codec should be used for the creation of preservation masters for digital

---

<sup>14</sup> NFSA, Op.cit.

<sup>15</sup> National Archives of Australia, Xena Preservation Software, *Supported formats*, available at: <http://xena.sourceforge.net/help.php?page=normformats.html>

<sup>16</sup> JISC Digital Media, *Choosing a digital audio file format*, op.cit.

video. Therefore, State Records cannot make specific recommendations regarding what particular creation formats should be used.

A preservation 'master' should be made using a suitable format of as high a quality as possible, then duplicating copies and reference copies can be made at a quality fit for purpose.

Your organisation should assess potential formats against the [Criteria for choosing formats for preservation purposes](#). See also [Formats and codecs for digital video preservation](#) for discussion about potential formats.

### **Choosing a format/codec for digitising analogue audio or video**

Please also refer to the advice given on [Digitisation of analogue audio and video](#).

State Records recommends that priorities for digitisation should include analogue recordings that:

- are masters and are used very frequently (as analogue recordings become degraded by use. Generally copies should not be made from masters)
- are already being degraded by high use and therefore in danger
- have the disposal action as 'retain in agency' or 'required as State archives' in the organisation's current retention and disposal authority (as these indicate the recordings that have significant ongoing interest to the organisation or the community).

#### **Short term**

Frequent use will not degrade digital materials (unless CDs or DVDs become scratched and damaged). Even if the file is a corporate record, if the digitised copy only needs to be kept for a very short period of time, or if high audio quality is unnecessary, your organisation should use the format and codec which is most suitable for the aims of the project and the editing software to be used. The processing power and storage space you have available may also impact on your decision.

Example: many of the MPEG file formats (e.g. mp3) which are standard definition, affordable and widely available may be used if the recordings are only required in the short term and fast delivery (e.g. over the web) is the goal. However, they use lossy compression and are not considered to be of sufficient quality for preservation.

#### **Long term or archival, or where quality is paramount**

If the audio or video file:

- will require long term or archival retention or
- needs to be of optimal quality

the format and codec for digitisation needs to be chosen very carefully. A preservation 'master' should be made using a suitable format of as high a quality as possible, then duplicating copies and reference copies can be made at a quality fit for purpose.

Digitising to a higher quality format (e.g. copying from VHS to Motion JPEG) will not increase the quality of the original analogue recording. However, the original recording should never be copied to a format of lesser quality when the aim is to create a preservation master. The aim should be to preserve as much quality of content as possible from the original recording. A lossless or uncompressed format will capture all of the original quality if the encoding to digital is optimal.

Formats should meet as many of the identified Criteria for choosing formats for preservation purposes as possible. However, sometimes these may need to be balanced with the resources available.

Only linear uncompressed codecs or mathematically reversible lossless codecs should be considered for preservation.<sup>17</sup> No compression is ideal, but the files may take up considerable storage space (particularly video files). Your organisation will also need the processing power to be able to open and view the files and do error checking etc.

### Audio

Widely recognised and accepted formats for preservation purposes for audio are WAV or BWF as these are uncompressed and generally use the lossless Linear Pulse Code Modulation (LPCM) codec.<sup>18</sup>

FLAC is a suitable format for digital preservation as it is stable, free and open source and employs lossless compression. The National Archives of Australia converts a number of other audio files (including WAV) to FLAC during the process of normalisation for storage in their digital archive.<sup>19</sup> If considering this format for business use, however, it is important to note that it is not widely supported by the audio industry.<sup>20</sup>

Note: Certain file formats may not be compatible with some audio software. You should also ensure compatibility with the software you intend to use for editing or optimising the file.

If the image is to be optimised, lossless methods should be used and information regarding the optimisation should be saved with the audio file separately from the preservation master. Other copies of the recording could use compression if necessary for delivery e.g. over the Internet.

For more information and issues to consider when choosing an audio file format see the JISC fact sheet Choosing a digital audio file format at <http://www.jiscdigitalmedia.ac.uk/audio/advice/choosing-a-digital-audio-file-format/>

See also Formats and codecs for digital audio preservation for more discussion about potential formats.

### Video

There is no consensus to date among the archival community as to which file format or codec should be used for the digitisation of analogue video which is required to be preserved in the long term or as archives. Therefore, State Records cannot make specific recommendations regarding what particular creation formats should be used.

Your organisation should assess potential formats against the Criteria for choosing formats for preservation purposes. See also Formats and codecs for digital video preservation for discussion about some potential formats.

Analogue video material is imported into a digital video format on a desktop computer by using an analogue-to-digital converter such as a video capture card. If digitisation is being done this way, the choice of capture formats may be limited by the small number of formats any given hardware capture card or

---

<sup>17</sup> NFSA, Op.cit.

<sup>18</sup> JISC Digital Media, *Choosing a digital audio file format*, op.cit., available at: <http://www.jiscdigitalmedia.ac.uk/audio/advice/choosing-a-digital-audio-file-format/>

<sup>19</sup> National Archives of Australia, Xena Preservation Software, *Supported formats*, available at: <http://xena.sourceforge.net/help.php?page=normformats.html>

<sup>20</sup> JISC Digital Media, *Choosing a digital audio file format*, op.cit.

software application can support. JISC recommend for mid-range desktop systems the losslessly compressed Blackmagic (proprietary) codec in an AVI or MOV wrapper or the losslessly compressed AJA (proprietary) codec in a MOV wrapper, which are widely used.<sup>21</sup> The National Film and Sound Archive suggest an uncompressed encoding in an MXF wrapper for small scale projects of this nature if the storage capacity is available. If not, MOV or AVI can be used with an uncompressed 8 or 10 bit encoding.<sup>22</sup> These may be converted to an open format later.

Several larger scale archival projects underway favour the open and lossless JPEG 2000 codec, e.g. with an MXF wrapper.

See also Formats and codecs for digital video preservation for discussion about potential formats.

### Criteria for choosing formats for preservation purposes

There are numerous formats for digital audio and video on the market. However, not all are suitable for preservation purposes.

It is recommended that if the digital audio and video records need to be retained and accessible in the long term or as archives, that your organisation chooses a format that is as sustainable as possible.

The Library of Congress has defined what they call 'sustainability factors' to look for. Not all formats will meet all of these criteria (or may meet them to lesser or greater degrees), but they provide a benchmark when choosing preservation formats.

- Disclosure i.e. Are the format's complete specifications and tools publicly and openly documented and accessible so that it is clear how the information is represented?
- Adoption i.e. Is the format in widespread use (a de facto standard)? If so, it is less likely to become obsolete and tools for replication are likely to be developed.
- Transparency i.e. Is the format open to direct analysis with basic tools such as human readability using a text-only editor?
- Self documentation i.e. Does the format contain basic descriptive, technical and administrative metadata? If metadata can be embedded the formats are likely to be sustained in the longer term (e.g. it can offer protection against centralised system failure) and transfer reliably from one system to another.
- External dependencies i.e. Does the format depend on particular hardware, operating systems or software for rendering or use? Ideally the format should be compatible with many software and hardware systems and made by different manufacturers so it will not be complex to manage in future technological environments.
- Impact of patents i.e. Is the format protected by patents? These may inhibit the ability to sustain content over time.

---

<sup>21</sup> JISC Digital Media, *Choosing a digital video file type*, 3 March 2009, available at: <http://www.jiscdigitalmedia.ac.uk/movingimages/advice/choosing-a-digital-video-file-type/>

<sup>22</sup> MOV is likely to be the better choice of the two as it is more widely adopted and has more metadata capability. Its proprietary nature is of concern, however. If this is unaffordable the NFSAs recommend, at the very least, well documented high-bit rate MPEG2 or DV encoding in the knowledge that there almost certainly will be a loss of quality.

- Technical protection mechanisms i.e. Are there mechanisms like encryption that prevent the preservation of content?<sup>23</sup>

JISC recommends consideration of some other practical factors to consider when choosing a preservation format including:

- Can it be repurposed?
- Is the format simple to use?
- Is the format evolving or stable?
- Can the format be kept secure and unalterable in a digital recordkeeping system?
- Is it expensive to use?<sup>24</sup>

Note: The sustainability of audio and digital video often depends on industry support, and there is often not commercial incentive to support some open formats. Industry driven formats, while often proprietary, may have publicly documented specifications and therefore may be useful for preservation. The danger is that proprietary formats are open to change by their owners that may affect the recoverability of the content. They may also have other drawbacks such as technological dependencies.

The Library of Congress has set up a [Digital formats website](http://www.digitalpreservation.gov/formats/) which provides information about digital formats which includes format descriptions and assessments against the above sustainability criteria for a range of sound and moving image formats (as well as other categories of digital resources). This is an excellent site to gain an understanding of evolving formats at <http://www.digitalpreservation.gov/formats/>

### Formats and codecs for digital audio preservation

A few examples of the container formats and codecs for audio that have been used for long term or archival preservation purposes are described below.

Note: As the availability and capabilities of formats and codecs change frequently and technological obsolescence is rapid, organisations embarking on digitisation or recording projects should check the relevant literature to determine which product best suits the project and preservation needs.

<b>The WAV family of formats</b>	<b>Advantages</b>	<b>Disadvantages</b>
Waveform Audio File Format (WAV) Developed by IBM and Microsoft. Based on the Resource Interchange File Format (RIFF).	Specifications fully documented and freely available Considered to be simple, easy to edit and manipulate with software International de-facto standard for digital audio:	Proprietary Inability to hold metadata describing its contents Uncompressed files are very large and are therefore unsuitable for some uses like

<sup>23</sup> Library of Congress, Sustainability of digital formats: Planning for Library of Congress collections, *Sustainability factors*, available at:

<http://www.digitalpreservation.gov/formats/sustain/sustain.shtml>

<sup>24</sup> J.Eaton, JISC Digital Media, 'Digital shelf life: building files to last', presented at *Decoding the Digital: A common language for preservation*, 27 July 2010, available at: <http://www.dpconline.org/events/decoding-the-digital-a-common-language-for-preservation.html>

	<p>in widespread use in the industry and accepted by the archival community</p> <p>Almost all audio players support WAV. Some normalise input formats into WAV before playing</p>	<p>transmission over the web and storage where there is limited space available</p> <p>Files are limited to 4GB in size<sup>25</sup></p>
<p>Broadcast Wave Format (BWF)</p> <p>Developed by IBM and Microsoft. Based on the Resource Interchange File Format (RIFF).</p>	<p>Specifications fully documented and freely available. See the European Broadcasting Union website (<a href="http://www.ebu.ch/">http://www.ebu.ch/</a>)</p> <p>Considered to be simple, easy to edit and manipulate with software</p> <p>Based on WAV which is international de-facto standard for digital audio</p> <p>Proposed as a preferred master format for mono or stereo audio and is optimal for broadcasters:</p> <p>Recommended by the International Association of Sound and Audiovisual Archives (IASA) in <i>TC 04 Guidelines on the production and preservation of digital audio objects</i></p> <p>Self-describing metadata</p>	<p>Proprietary</p> <p>Uncompressed files are very large and are therefore unsuitable for some uses like transmission over the web and storage where there is limited space available</p> <p>Files are limited to 4GB in size<sup>26</sup></p>
<p>Multichannel Broadcast Wave Format (MBWF)</p>	<p>Specifications fully documented and freely available. See the</p>	<p>Proprietary</p> <p>Uncompressed files</p>

<sup>25</sup> Information on WAV derived from Wikipedia, WAV, 11 July 2010, available at: <http://en.wikipedia.org/wiki/WAV>; JISC Digital Media, *Choosing a digital audio file format*, op.cit and *Uncompressed audio file formats*, 10 February 2010, available at: <http://www.jiscdigitalmedia.ac.uk/audio/advice/uncompressed-audio-file-formats/>; Library of Congress, Sustainability of digital formats: Planning for Library of Congress collections, *WAVE audio file format*, available at: <http://www.digitalpreservation.gov/formats/fdd/fdd000001.shtml>; Zytrax.com, *Digital audio/video formats*, available at: <http://www.zytrax.com/tech/audio/formats.html>.

<sup>26</sup> Information on BWF derived from Wikipedia, *Broadcast Wave Format*, 25 July 2010, available at: [http://en.wikipedia.org/wiki/Broadcast\\_Wave\\_Format](http://en.wikipedia.org/wiki/Broadcast_Wave_Format); Library of Congress, Sustainability of digital formats: Planning for Library of Congress collections, *Broadcast WAVE audio file format*, available at: <http://www.digitalpreservation.gov/formats/fdd/fdd000003.shtml>.

<sup>27</sup> Information on MBWF derived from Wikipedia, RF6, 21 September 2010, available at: <http://en.wikipedia.org/wiki/RF64>

Developed by IBM and Microsoft. Based on the Resource Interchange File Format (RIFF).

European Broadcasting Union website (<http://www.ebu.ch/>)

Considered to be simple, easy to edit and manipulate with software

Based on WAV which is international de-facto standard for digital audio

Recommended for use for surround sound or multichannel audio

Self-describing metadata

Files not limited to 4GB (restrictions on WAV and BWF)<sup>27</sup>

are very large and are therefore unsuitable for some uses like transmission over the web and storage where there is limited space available

It is generally believed that the WAV family of formats will resist obsolescence and when they do become obsolete that the tools to manage their replication to other formats will be available and informed by the professional audio industry.

### Free format Lossless Audio Codec (FLAC)

Codec developed by Xiph.org a non-profit organization. There is a simple container for the stream also known as FLAC.

### Advantages

Free, open source and patent-free

Self-documenting: metadata is contained in the file header

Uses completely lossless compression so when decompressed there is no loss of quality

Uses less disk space than WAV files containing the same content, while maintaining full audio quality

A number of tools can be used to write and render the format

Frequently used in high quality playback systems

### Disadvantages

Poorly supported by playback and editing software and hardware e.g. iTunes for Mac OS and iPods will not play FLAC format

Without major support from the audio industry, may become obsolete more quickly<sup>28</sup>

<sup>28</sup> Information on FLAC derived from Wikipedia, *Free Lossless Audio Codec*, 2 August 2010, available at: [http://en.wikipedia.org/wiki/Free\\_Lossless\\_Audio\\_Codec](http://en.wikipedia.org/wiki/Free_Lossless_Audio_Codec); Library of Congress, Sustainability of digital formats: Planning for Library of Congress collections, *FLAC (Free Lossless Audio Codec) Version 1.1.2*, available at: <http://www.digitalpreservation.gov/formats/fdd/fdd000198.shtml>; JISC Digital Media, *Choosing a digital audio file format*, Op.cit.

	and some archival applications	
<b>Audio Interchange File Format (AIFF)</b>	<b>Advantages</b>	<b>Disadvantages</b>
An uncompressed audio file format developed by Apple Computer for storing sound data for PCs and electronic audio devices.	<p>A leading format used by professional-level audio and video applications</p> <p>Uncompressed (which aids rapid streaming of multiple audio files from disk to the application) and lossless</p> <p>There is a compressed variant of AIFF known as AIFF-C which can use a "pseudo-compressed" codec called sowl. The data can be converted back and forth without loss</p> <p>Self documenting metadata: each chunk has metadata contained in the file header</p> <p>Widely adopted among Apple users and some other users</p>	<p>Most Windows users select WAV, which is functionally equivalent</p> <p>Uncompressed files are very large - not suitable for some uses like transmission over the Web and storage where there is limited space available<sup>29</sup></p>

See the Library of Congress' [Digital formats website](#) or Wikipedia for a summary of each format. at

[http://www.digitalpreservation.gov/formats/fdd/browse\\_list.shtml](http://www.digitalpreservation.gov/formats/fdd/browse_list.shtml)

### Other requirements

A general rule that recording is done at a minimum sample rate of 48KHz (96kHz is considered optimal for BWF) and 24-bits (depending on the type of recording).<sup>30</sup>

### Audio files not recommended for preservation

While they may be useful for short term records, or to meet particular needs, State Records does not consider the following digital audio files to have sufficient quality for preservation purposes (although they may be used for short term records or copies if required):

- Files created for streaming broadcast (e.g. RealAudio, Windows Media)

<sup>29</sup> Information on AIFF derived from Wikipedia, *Audio Interchange File Format*, 2 August 2010, available at: [http://en.wikipedia.org/wiki/Audio\\_Interchange\\_File\\_Format](http://en.wikipedia.org/wiki/Audio_Interchange_File_Format) ; Library of Congress, Sustainability of digital formats: Planning for Library of Congress collections, *AIFF (Audio Interchange File Format)*, available at:

<http://www.digitalpreservation.gov/formats/fdd/fdd000005.shtml>

<sup>30</sup> NARA, *Ibid.*; JISC *Choosing a digital audio file format*, Op.cit.

- Reference files of lower quality than the original (e.g. for use on the website)
- Most audio files in the MPEG format (e.g. mp3).

These are generally too compressed, sacrificing quality for file size.<sup>31</sup>

Note: If an organisation has a master in a compressed format it would be considered the preservation version by default. A lossless clone could be created as a WAV or FLAC file so that it can, in future, be migrated without further loss of quality.

### Formats and codecs and formats for digital video preservation

There is no consensus to date among the archival community as to which file format or codecs should be used for preservation purposes for digital video. Therefore State Records is unable to recommend one particular format over another at this time.

Some formats that are used for preservation include:

Format name	Description
Audio-Video Interleave format (AVI)	A multimedia file format developed by Microsoft and IBM. A derivative of the RIFF format. It is well understood by many editing and transcoding systems. This is now virtually obsolete and has limited metadata capacity but is widely supported.
Material Exchange Format (MXF)	An object based file format developed by the Society of Motion Picture and Television Engineers (SMPTE) that wraps video, audio and other bitstreams. The MXF standard has been increasingly adopted in broadcasting. See the European Broadcasting Union website ( <a href="http://www.ebu.ch/">http://www.ebu.ch/</a> ) for technical reviews.
QuickTime format (MOV)	A multimedia file format developed by Apple Computer Inc that wraps audio, video, effect or text. It is well understood by many editing and transcoding systems. Note: While widely supported with good metadata capacity, this is proprietary.
Motion JPEG 2000 file format	An object oriented file wrapper developed by the Motion Pictures Experts Group (MPEG). It specifies the use of JPEG 2000 for timed sequences of images (motion sequences) possibly combined with audio.
JPEG 2000	A still image codec created by the Joint Photographic Experts Group in 2000 and published as an ISO standard, ISO/IEC 15444. JPEG2000 has been increasingly adopted in moving pictures sectors.

---

<sup>31</sup> NARA, Ibid.

DIRAC	An open, royalty free video compression format, specification and system developed by the BBC.
FF video codec 1 (FFV1)	A lossless, intra-frame video format. The encoder/decoder is part of the free, open-source library libavcodec in the project FFmpeg. <sup>32</sup>

Only linear uncompressed codecs or mathematically reversible lossless codecs should be used for preservation.<sup>33</sup>

For more information on formats, including their sustainability, you could start by examining their profiles on the Library of Congress [Digital formats website](http://www.digitalpreservation.gov/formats/) at <http://www.digitalpreservation.gov/formats/> or their individual descriptions on Wikipedia. See also the JISC fact sheet *Choosing a digital video file type* at <http://www.jiscdigitalmedia.ac.uk/movingimages/advice/choosing-a-digital-video-file-type/>

Note: Work is underway in the US to establish a permanent archival standard using lossless JPEG 2000 with an MXF wrapper for all forms of moving image archival encoding.<sup>34</sup>

The Library of Congress' own National Audio-Visual Conservation Center, is implementing systems that reformat older videotapes as lossless JPEG 2000 frames wrapped in MXF, i.e., MXF SAMMA.<sup>35</sup>

The National Film and Sound Archive, Australia has announced that it has also purchased SAMMA digitisation systems to use JPEG 2000 frame images, losslessly compressed with an MXF wrapper for the digitisation of more than 40,000 hours of video. See [Case studies of digitisation of analogue audio and video](#) for more information.

The National Archives of Australia are also converting their analogue video to MXF wrapped JPEG2000.

Note: The above solutions are for large scale projects. At present it is not easy to use standard video workstations for the efficient preservation of digital video content. For example, the National Film and Sound Archive advise that almost all workstations on the market do not provide capture to lossless JPEG2000 and can only provide archival acceptable formats that are uncompressed.

If organisations have a small collection of valuable video material, they may need to contact an outsourced provider to digitise them. [Appendix 2: Outsourcing the digitisation of audio or video](#) includes a list of considerations when using outsourced providers.

### Other requirements

When creating digital video or digitising analogue formats, it is generally recommended that:

- the audio tracks of digital video recordings should be at a minimum of 48KHz (the broadcast sample rate is always 48 kHz and 16, 20 or 24 bits depending on the age of the format and equipment)

<sup>32</sup> Information derived from Wikipedia and the Library of Congress Digital formats website at <http://www.digitalpreservation.gov/formats/>

<sup>33</sup> NFSA, Op.cit.

<sup>34</sup> NFSA, Ibid.

<sup>35</sup> Library of Congress, Sustainability of digital formats: Planning for Library of Congress collections, *Material eXchange Format (MXF)*, available at: <http://www.digitalpreservation.gov/formats/fdd/fdd000013.shtml>

- a minimum of 720x486 pixels at 25 frames per second is achieved
- the colour depth match the number of colours as well as the colour encoding and luminance of the original material. Black and white originals should be recorded in grayscale; colour originals should be recorded in full-range colour (8 bits per channel with 10 bits per channel preferred).<sup>36</sup>

### Formats that do not have sufficient quality for archival retention

While they may be useful for short term records, or to meet particular needs, the following digital audio files do not have sufficient quality for preservation purposes (although they may be used for short term records or copies if required):

- Files created for streaming broadcast (e.g. RealAudio, Windows Media) as they are too compressed and sacrifice quality
- Reference files of lower quality than the original (e.g. for use on the website) as they are generally too compressed and sacrifice quality
- Files that have been through two or more codecs.<sup>37</sup>

### Quick tips regarding digital file formats

- Choose a format that is suitable to the project aims and the longevity of the recording.

If formats are required in the long term or as archives:

- Use the best quality format you can afford for creation and/or digitisation.
- Use no compression (or lossless compression if compression is unavoidable) for creating a digital preservation master. Lossy compression irretrievably reduces quality.
- Use an open format or a format which has good 'sustainability factors' for the preservation master. If you can't create in such a format, copy to one when possible.
- Define your metadata requirements carefully and ensure metadata is collected from creation.
- If enhancements are to be made, make these from a copy. Save a 'master' of the original and enhanced version.
- Create copies for delivery/access in formats that meet your needs (e.g. lower resolution formats).
- If possible, store a copy of your preservation master (and enhanced master, if any) at a different location for disaster management purposes.
- Regularly monitor your masters for obsolescence or degradation.

### Case study: Tourism NSW - Use of digital video to promote the State

Tourism NSW is a branch of the Department of Industry and Investment NSW. It aims to increase visitor numbers to NSW and contribute to the economic growth of NSW businesses and the NSW community.

<sup>36</sup> NARA, Op.cit. In general component colour is used Y/U/V. There are many standards for encoding YUV 4:4:4, 4:2:2, 4:2:0 and all affect the quality. 4:4:4 is the ultimate, but as there is too much data in a 4:4:4 encoding, most archives use 4:2:2.

<sup>37</sup> NARA, Ibid.

The growing availability of radio, television and the Internet in recent decades has meant that audiovisual materials are essential tools used by NSW Tourism to showcase the State.

Tourism NSW has a video collection of campaign and promotional material that dates back to the late 1980s. Most of this collection is on Digibeta (Digital Betacam) with viewing copies on VHS tapes. As audio-visual materials date quickly much of this collection is not used for current campaigns and promotions, but may be used for anniversaries and celebrations. For example, footage was used in the celebrations of 100 years of tourism in 2005. It is also requested by the media, students and other individuals. Tourism NSW is currently arranging the transfer of some of this collection to State Records as they are required as State archives in its retention and disposal authority.

Current campaign and promotional materials are made by external contractors who maintain the masters in a high quality format to enable reuse. Tourism NSW specify in their contracts that the material needs to be supplied to them in suitable formats, which is usually on Digibeta and DVD. The copy in Digibeta is still considered to be the internal 'master' and the DVDs are for delivery to internal and external users. Current footage is often requested by the media, other tourism bodies and tour operators (e.g. for use on their websites, in buses etc). Tourism NSW currently makes the footage available at no cost as its use will help to promote the State.

Increasingly Tourism NSW is also featuring footage on its website or making it available on other Internet sites. Generally this footage will still be shot in a high quality format to enable reuse but will then be copied into a lower resolution format for online delivery.

Tourism NSW has identified the value of having detailed metadata for their footage in order to meet their business needs. They create 'cue sheets': spreadsheets that detail every shot of the footage and its location (this can also promote reuse of the footage). They also create 'call sheets' which include information on the actual shoot e.g. day, location, time, the crew and the talent involved. Contracts are signed between members of the case and the production crew which detail usage and copyright issues. Further metadata including unique identifier, description of contents, format, usage restrictions, date created etc is added for internal control of the tapes.

### **Digital betacam**

Digibeta (Digital betacam) has been in common use since the 1990s and for many years was considered a sound archival medium. Although the equipment to read Digibeta is still widely available it is becoming less popular now as it is tape-based. This prevents automated error checking and it is difficult to give immediate access to the footage.<sup>38</sup> Tourism NSW has faced these issues and is now looking at the viability of a digital hard disc based storage system to take advantage of the benefits of digital formats.

---

© State of New South Wales through the State Records Authority, February 2009. Part 5.5 added 2010. This work may be freely reproduced and distributed for

---

<sup>38</sup> JISC Digital Media, *Transcoding digital video*, 27 January 2009, available at: <http://www.jiscdigitalmedia.ac.uk/movingimages/advice/transcoding-digital-video/>

most purposes, however some restrictions apply. See the copyright notice on [www.records.nsw.gov.au](http://www.records.nsw.gov.au) or contact State Records.

ISBN: 978-0-9805148-9-6