BROADCASTER'S DILEMMA WITH ARCHIVE ASSET MANAGEMENT: TORN BETWEEN LONG TERM AND PRODUCTION REQUIREMENTS

Silvester Stöger, NOA, Austria
Jean-Christophe Kummer, NOA, Austria

1. Broadcaster's Dilemma

The media industry was shaped by fundamental changes in the last decades. Contemporary capturing, editing, and production, as well as distribution, have become digital and file based. In contrast to this reality, institutions such as broadcasters, distributors, or cultural heritage organizations have accumulated hundreds of thousands of hours of content bound to physical carriers, both analogue and digital. Those tapes, records, films, cassettes, disks, or drives are sitting in the archives and can be of great value. However, due to obsolescence of players and physical deterioration of the legacy media, recordings may become unplayable and the content they were meant to preserve can be lost.

As a consequence, broadcasters often find themselves faced with a dilemma as they strive to manage their current production and distribution on one hand and preserve their legacy and contemporary productions on the other. Within this, the setup of file-based repositories is an essential strategy, because obviously content can be made both easily accessible and readily available for use in production and distribution.

The setup of a digital archive poses new challenges, especially for the conservation of investments. Whereas millions might have been spent to create an hour of content, the many thousands of hours lodged in an archive are often considered to be a 'free of charge' asset which requires no further expenditure.

The broadcaster's requirements can be summarised as:

- **Hundreds of thousands of hours of archive content:** The broadcaster is interested in not losing valuable archive content, but the sheer size of the job poses fundamental difficulties in preservation.
- **Long term preservation of archive content:** Archive content needs to be there for generations to come, as historical artefacts are a proof of history and possibly also monetisable resources.
- **Obsolescence of players / Deterioration of carriers:** The progress of technology and the nature of physical carriers (aka as "degralescence[1]") make continued preservation on the original media unviable.
- **Migration of content bound from physical carriers to the digital domain:** Migration to the digital file based domain will be the only option, but various issues related to physical carrier and digital file formats have to be overcome during migration.
- **Choice of archive format:** What file formats fulfill all archival requirements? Will I prefer a multi-format archive or a normalised archive format? Will I select a dedicated archival format which is possibly lossless and not lossy as usually in production?
- **Feed archive content to production system:** Archive content is valuable if it can be reused in production.

1 Mike Casey, "Degralescence: the combination of obsolence of replayers with degradation of carriers," IASA journal, no.44, (January 2015): 17.

- **File based production:** All production workflows are now file based, from recording to distribution.
- **Periodically changing production formats and environments**: Production file formats as well as production systems change with the market introduction of new standards and technology.
- **Distribution of content to various platforms:** In the 21st century a broadcaster needs to play out content to more than one distribution channel. The traditional TV and radio channels have to compete increasingly with various Internet distribution platforms, including in-house and external on-demand services.
- **Dissemination platforms change rapidly:** As the media market grows more competitive, a broadcaster will have to become more flexible and adapt to new forms of distribution. Production, Archive and Distribution should therefore be streamlined while staying compatible.

The requirements above point to the growing importance of a **repository** which specifically takes care of archival data and which could be described as:

- Stores less frequently used content (think about production footage from the 1960s)
- Stores legacy content from the historic beginning of a TV station until the file based era
- Requires the ability to maintain interfaces to many different kinds of production systems
- Enables a broadcaster to quickly overcome its degralescence problem of carrier based audio and video archives in an easy to manage new home with enhanced descriptive metadata
- Tracks the description of and the link to the legacy physical formats – specifically in transition phases (co-usage of physical material and digitised material is important, e.g., in film collections)

For this paper, let´s call this new repository, **Archive Asset Management (AAM)**.

Such a repository only becomes a secure and independent archive if the architecture obeys OAIS[2] principles of archive functions and contains rich description and metadata. Archive assets (media essence, and metadata with descriptive rights) need to be taken care of with a special ethical framework that may be different from the norms of production environments, while ensuring full performance in daily business.

2. Archive Asset Management (AAM)

2.1 Archive assets

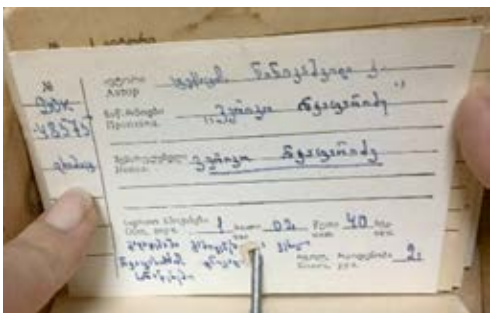
2.1.1. Essence

All archive essence can be digital data nowadays and any content on physical carriers will very likely need to be migrated to the digital domain in order to be disseminated. A digital archive offers many possibilities, but to preserve all investments, good care must be taken in the design of the system. If essence is to remain in good condition, a thorough Archive Asset Management (AAM) is necessary. An AAM may take care of media management and integrity, enables extensive metadata annotation, executes all transactions within the archive, and serves as a bridge to attached production systems.

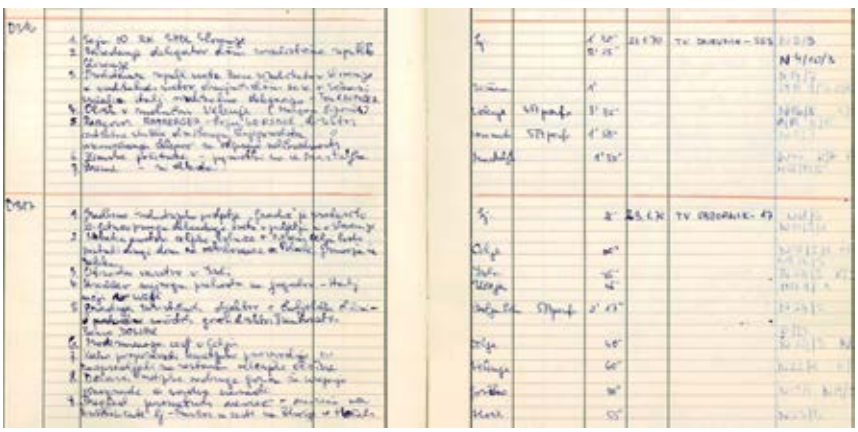
2 "ISO 14721:2012," Standards catalogue, International Organization for Standardization, accessed September 29, 2017, https://www.iso.org/standard/57284.html.

2.1.2. Metadata (and rights, and essence) = asset

Archive Assets are not just essence files, they get their value from descriptive metadata. Metadata is the key to unlock and access content inside an archive. However, a production system may not allow extensive enrichment of content with descriptive metadata. Only basic information useful for production is intellectually created in production environments, besides in some cases automatic technical metadata (subtitling, or technical information such as geotags or EXIF data). Although this may be sufficient for production workflows, to search the archive for unique archive footage or rare material, technical metadata and basic descriptive comments, especially in the context of legacy carrier-born data, do require further annotation with structured and customizable descriptive metadata. Only if media gets enriched with descriptive annotation, files will become content that is searchable and useable. A management system that does not allow complex annotation of content cannot be considered a practical archive, but rather a large and very expensive black box. Even if cloud based tagging tools help enrich content, a basic description of, for example, a Georgian field recording or Slovenian metadata from a broadcaster need the intellectual resources of an archivist to make it clear, especially for legacy content, what the recording contains.

Example: Metadata Card Georgian Broadcaster – handwritten (Georgian font)	Example: UAE – National Broadcaster (handwritten Arabic language)
	

Example: Metadata Card Slovenian National Broadcaster (handwritten Slovenian language)



It turned out that some unsuccessful projects were reported[3] where outsourcing of metadata annotation to an external service provider caused more problems than gain, requiring additional remediation after the fact. This gets more obvious when looking at the source of metadata above which is far from being easily digitally processable.

2.2 Archive migration

Traditionally archives are the heart of institutions. Over time, essence has been recorded on physical carriers, collected in the archive, and kept safe there to survive the course of time. While broadcasters focus most attention on daily production and distribution, the archive is accessed when content is requested for reuse. These are the moments when an archive proves its value. Archive footage has not only cultural importance as historical documents (see also "UNESCO World Day for Audiovisual Heritage"), but also demonstrates its economic merits by being reused in production of new programs. For a broadcaster, those two concepts are tightly linked.

Figure 1. Archives that do not exist in this shape anymore: left: digitized obsolete carriers at ORF, right: Syrian archive destroyed in war.

Producing fresh content for television or radio is always linked to costs, just like archive footage was, when it was originally produced and edited. Keeping tapes safely in the archive comes with a price as well (e.g., air conditioning, space rental and maintenance, and player maintenance), but at the end of the day all the cultural and economic value stored in a broadcaster's archive is worthy of preservation—for the present and the future. It would be grossly negligent not to sustain an archive and its holdings or to allow the treasures contained to degrade, especially when the small fortune it cost to produce is taken into account.

2.2.1. Mass migration of physical archives

Having considered the threats to archives and the needs of the holding institutions, mass digitisation has turned out to be the obvious precondition for future survival of legacy archive content. This way all carriers in the physical archive will be transcribed into a digital file based repository (as virtual copies of the physical original). This should happen in a planned manner and with quality controlled processes. Of major importance is the well-orchestrated ingest of digitized carriers into an archive storage as well as writing an index of the created files into a central database. One of the problems broadcasters face, is that existing repositories used in Production Asset Management (PAM) tend to handle only file based information, ignoring physical existence of legacy archives.

3 Ilse Assmann, Quoting her experience with metadata outsourcing at M-Net / South Africa, Broadcast Section reports during IASA conference Washington, September 2016.

Further PAM´s are mostly optimized for specific content (i.e., film, or audio, or video) and support only a very specific production file format, which might be very likely not suitable for long term archiving (such as a lossless format), resulting in various PAM's inside one institution. Each of those production systems is equipped with its own storage and file management layer. So the need of a central management layer arises.

Industry has answered this need by extending PAM´s with additional Media Asset Management (MAM) systems—possibly addressing the need of archival descriptions and finding a way to help archives in both worlds, by copying database indexes to a kind of umbrella MAM and insisting on a jack-of-all-trades system. Typically MAM and PAM systems are tightly integrated and most compatible if provided by one vendor only. This makes the exchange of single systems quite difficult, and meets the interests of manufacturers.

The description of multi-carriers and pointing towards still existing archive carriers is only handled via some descriptive information via prefixed ID´s—in the case where this at all possible.

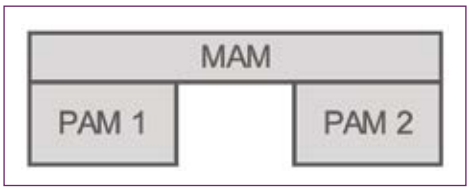


Figure 2. Tightly integrated PAM's to an umbrella MAM.

Facing the fact that legacy archive content should be archived for the long term, it is very questionable if digitized content should be ingested to such a tightly-knit systems cluster. Given the situation of changing PAMs and changing dissemination paths, a strategy can be to think of a lightweight asset management system, which does NOT include all production elements of editing and playout, but concentrates its management functions on the dedicated archive requirements. This lightweight and specialized Archive Asset Management system (AAM) resolves the stigma of abusing an existing PAM which might not be useable for the transition of physical archives to file based content.

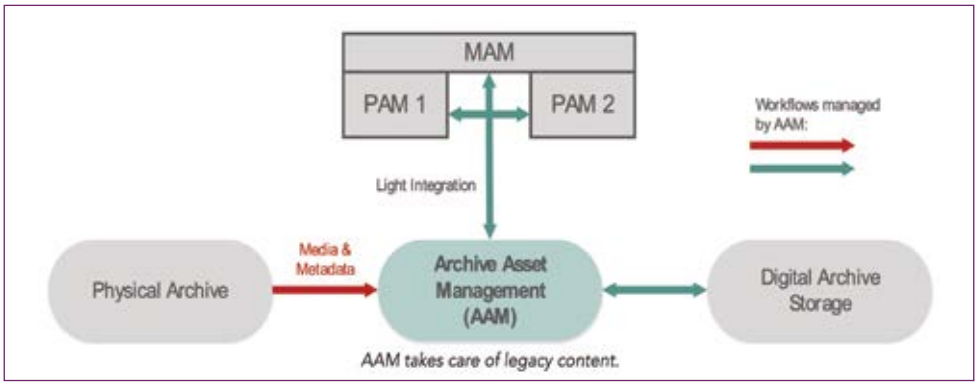


Figure 3. AAM takes care of legacy content.

2.3 Characteristics of specialized Archival Asset Management (AAM)

2.3.1. AAM ≠ PAM

The differences between an archiving and a production system cannot be stressed enough. Production Asset Management (PAM) systems seldomly fulfil all archiving requirements. Regarding the management of the various sources and different types of metadata, systems specialized in production will rarely meet the requirements.[4] It should be verified if they follow other fundamental archival principles as it is a common misperception to understand production asset management equally as archiving. The long-term preservation of digital assets within an archive asset management framework should enable the conservation of information following general archival guidelines. In the context of a broadcasting institution with different fields of activity (production, dissemination and archiving) it is difficult to find a jack-of-all-trades platform for all the stakeholders. Typically large MAM systems claim to fulfil these needs and have been developed in the past under the context of gluing many components together.[5] For archive management, it should be therefore considered that a specialized system, designed for the longevity of archive content could be a solution which allows for handling preservation at a higher speed than trying to solve all MAM problems of an institution in one go.

2.3.2. Specialized archive backend

An Archive Asset Management (AAM) system is a specialized archive backend that has to provide longevity to the content inside the digital store—and, some say that "longevity in an archival context is eternity plus one day."[6]

A well-designed AAM could therefore easily outlive the attached production environment with its tightly knit additional Media Asset Management (MAM), Production Asset Management (PAM) system, editing suites, playout systems, and possibly proprietary storage.

Those components are configured for best performance, are vendor specific, and highly customised. Moreover, they might have been set up for current media standards and will be replaced on a regular basis when production and distribution standards change. So that the regular exchange of components does not affect the stability of the digital archive, the AAM should be a separate system, unaffected by the production environment but coexisting with it and its associated technology. In order to enable the continual life cycle of an AAM system, it needs to run on standard IT equipment that can be updated and exchanged easily if the necessity arises. The AAM must not be locked to any vendor-specific hardware but be prepared for change and future development.

Furthermore, the prediction of Moore´s law requires the usage of an independent storage system which can be easily exchanged once new technologies arise. This is a major difference from many systems deployed in the storage and production domains.

4 "Canalizing the maelstrom of metadata: extensions on the hourglass model," Brecht Declercq (2016) http://www.den.nl/art/uploads/files/DECLERCQ%20Canalizing%20the%20maelstrom%20of%20metadata%20(2).pdf, last accessed 22.11.2017

5 TecmathAG Kaiserlauten (2000), then BlueOrder (2004),– significant milestones in MAM development happened in the YLE Metro project around 2007 as bridge between PAM, legacy TV databases and playout systems, later it has been acquired by AVID MAM (acquisition in 2010).

6 Peter Bubestinger, JTS Singapore, March 8, 2016.

The infrastructure situation at a broadcaster can be summed up as a wide range of different specialized systems, all fitting their purposes. The specialized AAM, in turn, would be the piece of the puzzle taking care of preservation and description of legacy content.

2.3.3. Serving production and distribution environments

Constant availability of archival footage in digital file-based form is a major performance factor. After digitisation, the entire workflow of ordering content from the archive is no longer dependent on physical legacy carriers and can be handled in the AAM by automatic business procedures without unproductive waiting times. Serving distribution channels quickly with unique archive content is a major asset for broadcasters or cultural heritage institutions. It can then be searched, browsed, and previewed in the central Archive Asset Management system. Exports from the archive system to the production system with live transcoding need to be processed with minimal delays. If big repositories use tape based storage such as LTO tape libraries, the storage systems have to be scaled according to the expected access requests so that no queue will build up. The automatic retrieval of content from tape storage technology and transcoding are the only waiting times in feeding content from the archive to production. The readiness to deliver to different existing and future channels may be fulfilled by an independent archive system, agnostic to playout destinations and future file formats.

In addition to as yet unknown platforms, current distribution channels could include:

- Internet Radio, Internet TV, Video-on-demand
- Asset Selling Platforms to re-monetize content
- New platforms serving university networks or schools
- Platforms aimed at content collaboration

2.3.4. Coexistence with other systems and separated access domains

OAIS is a conceptual framework for how to design a secure and future-proof digital archive. It describes how the system architecture needs to have transparent and separated domains for producers of content, the actual archive system, and the distribution channel. In order to follow this design, an AAM system and the peripheral infrastructure should be structured in separated access domains. The AAM itself is located in the Archive Domain. This is traditionally the same physical department as the legacy archive or the digital archive storage, and any access transaction is managed by the AAM utilizing preconfigured workflows. The production and distribution systems themselves are separate domains. Any access transactions between AAM and PAM / MAM systems are also managed utilizing preconfigured workflows, mostly by the AAM, depending on the implementation design of the systems. Any access to the archive, such as search queries, media previews, or orders of content are triggered from user interfaces that are located in the user domain. Access procedures are also managed by the AAM and should be restricted with user rights management according to access permissions. By designing a digital broadcast archive in this manner, each element in the infrastructure stays independent, meaning it can be replaced if necessary without affecting the entire infrastructure; only communication between the systems has to be reconfigured. Further, the AAM ensures that the archive stays consistent and secure. No unauthorized access to archive essence is permitted and different access levels and permissions to alter the archive content can be managed with user rights. All transactions and manipulations inside the archive are executed by preconfigured workflows, this way every change is based on the same business processes, resulting in uniform archive content.

Digital Archive Storage in use has to be migrated or even changed in a 3–5 year period to overcome obsolescing technology—having an open interface to HSM or general storage systems helps you to fulfil this task.

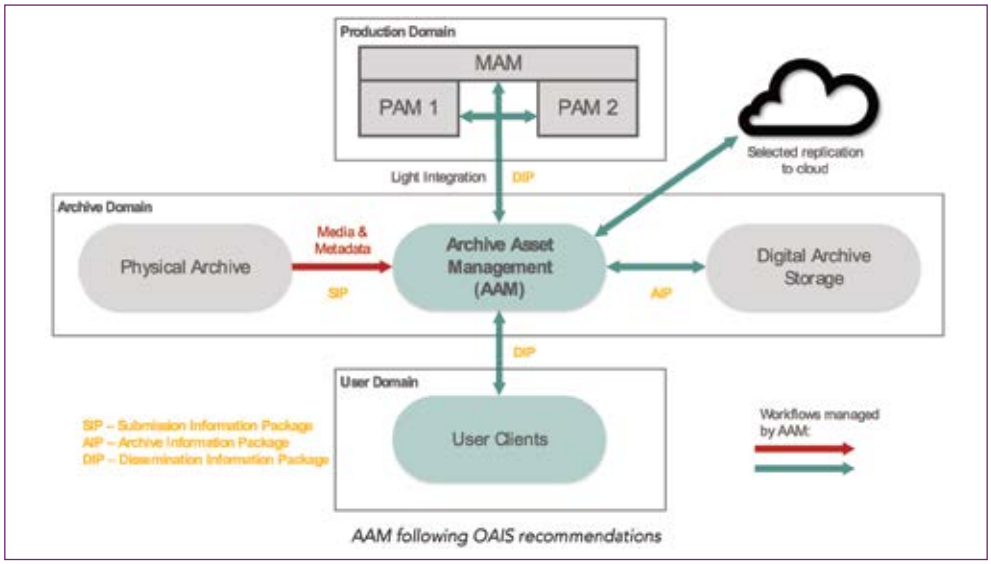


Figure 4. AAM following OAIS recommendations.

3. Essential AAM Elements

Archive Asset Management should be built on three main pillars: media, metadata and workflows.

Media	Metadata	Workflows
■ Archive essence data ■ Ensures best practice in media management	■ Description of media (essence) data ■ Metadata is key to finding and accessing archive content ■ Metadata covers rights information ■ Archive: black box or well- structured catalogue?	■ Processing of information ■ Ensures consistency of archive and business procedures ■ Provides workflow management

3.1 Media

Media is understood to be the archive essence data, the actual manifestation of archive content. To fulfil the purpose of an AAM, the system should ensure best practice media management. The purpose of media is that media files become content through our ability to access them.

3.1.1. Parallel management of multiple representations of media

Archive Asset Management does not just organize archive essence media. For the usage of the archive system additional media management features need to exist. One of them is the handling of multiple representations of media. A typical scenario is that media exists in the archive file format used for long term preservation of content, but high-resolution files are not practical for preview (streaming over network and retrieving from the archive storage), so preview proxies, a low resolution copy of the same content are typically provided. Potentially a third version of media has to be managed in an AAM. In NOA´s mediARC systems, for example, this is called auxiliary media, characterized as supplementary versions of the same media. An example could be a common distribution format likely to be used again (e.g., DCP in cinema distribution) or any original file format in which the content was produced (in this case it could be important to keep the original if any significant properties are bound to the original file format). Ideally, internal working formats that use mathematically lossless compression and archival formats that use mathematically lossless compression will be chosen, so no significant properties should get lost in transcoding digital media to an archival format.

3.1.2. Normalisation to one file format

An important approach when building a new archive is to avoid multiple archival master formats. Looking into today's archive shelves, you may find around 20 different physical carrier formats and each one of them requires specific knowledge, specific equipment, and specific treatment. In a file based digital archive, formats that are agnostic to content can be chosen. Some fundamental rules to guide the decision are: linearity, lossless compression, and safeguarding of the original's significant properties such as resolution or colour space.

If an archive is designed with a normalized format that uses mathematically lossless compression, it will be much easier in 15 to 20 years, to migrate from one archive format to another, rather than from twenty different formats. Additionally, streamlined business processes without generation loss are possible when using a lossless compressed format. This results in a transcoding history without generation loss.

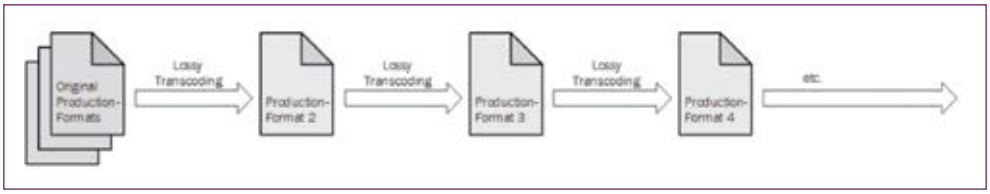


Figure 5. Transcoding history with generation loss.

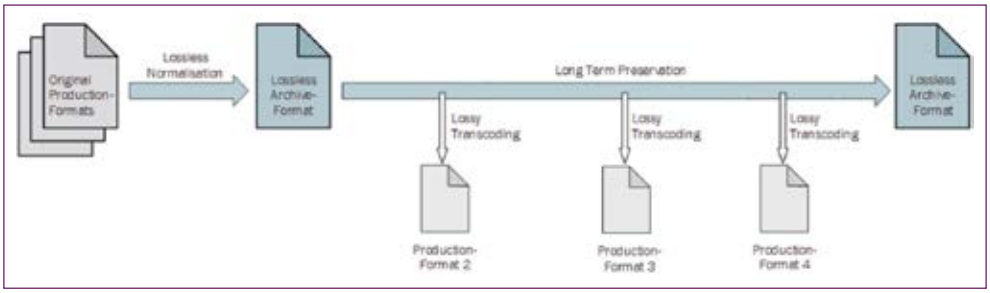


Figure 6. Transcoding history without generation loss.

3.1.3. Archive file format

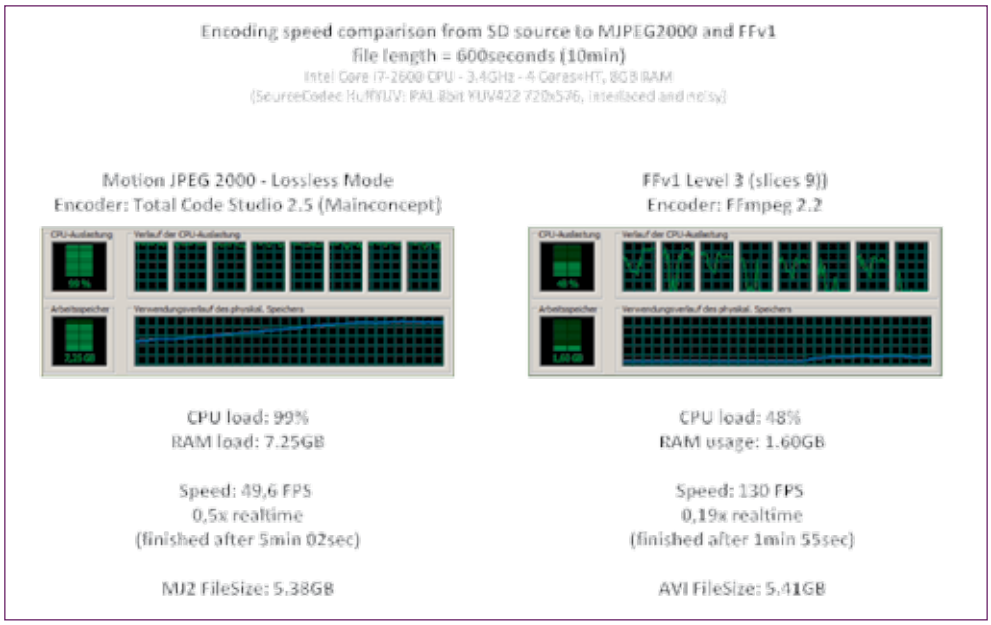
In the context of digital archiving, the choice of the correct video file format for legacy archiving has been one of the predominant topics of discussion for a long time. A few concepts should be considered to guarantee long term preservation of media. The most important is that archival formats that use mathematically lossless compression should be facilitated in an Archive Asset Management system. Often the argument is brought forward that a mathematically lossless archive format will raise storage costs (as opposed to using lossy compression formats). But storage cost is decreasing continuously. The fact that the radio archives community had intense arguments for 256kbit/sec encoding formats in 1995 because of storage costs seems ludicrous today (and arguing for DV25 will seem similarly strange in 2026). The time required to digitize material (most projects have durations of several years) will penalise any decision that optimises according to storage cost. The minor increase of storage costs is well worth the advantage of not reducing quality during ingest. Furthermore, the archival format constitutes a real essence master, as it represents the original content in its most authentic quality and allows future format migration without introducing further quality loss. Additionally, it is good practice to build an archive with one media format per media type; this reduces operational intervention during future migration, making it faster, cheaper, and simpler to maintain than multi-format archives.

Audio: A common ground has been established for audio archiving that is based on 48–96 kHz / 24bit PCM samples stored in a RIFF-based Broadcast Wave File. This will be adequate for both uncompressed originals and currently compressed media. However, in the cases of currently compressed media, associated metadata (e.g., from MP3 tags or MP2 information) will need to be retrieved and kept as separate metadata.

Video: The situation is more complicated here. We see the following situation in archives when it comes to selecting a normalised file format for SD material:

Archives with high demands on materials (e.g., the US Library of Congress)	**Broadcasters (e.g., ORF, YLE)**	**New Approach (e.g., RTVSLO, Slovak TV, Sharjah Media Corporation)**
Lossless format (JPEG 2000)	-	Lossless Format (FFV1)
Mezzanine Format (DVCPRO50)	Mezzanine Format (IMX50, DVCPRO50)	-
Preview Format	Preview Format	Preview Format

When asking why we need a mezzanine format, in the first case, it gets clear that the available software implementations of JPEG 2000 seem to be clumsy if transcoding times from archive to production take unnecessary long:



We have seen that FFV1 has gained traction against JPEG 2000, perhaps because of the situation below (as of December 2017):

	FFV1	JPEG 2000
Mathematically lossless?	yes	yes
Container	MKV (Matroska)	MXF (Material Exchange Format)
Pacemakers	Internet Community & Archivists	Commercial Companies & FADGI
Normative Institution	IETF	SMPTE
Specification	Cellar	AS-07
Disclose working Source Code of Encode Technology?	Yes (Open Source - FFMPEG)	No (Kakadu, Comprimato)
Standard ready?	no	no
Working pre-versions	Yes – 16 years – open source	EVS – yes for 4 years (1 company)
Computing Power	low	high
Typically suitable for? ($$$)	Small Budget Archives to Large Institutions	Large Institutions

By encouraging the open FFV1 video format (currently in the process of IETF standardisation), a future-proof encoding format to serve the purpose of building long-term video repositories may be found. It also might be interesting to see that adoption has taken place even with the standard not being ready, as working (open source) implementations exist.

Broadcasters still tend to have a short-term perspective and often decide not to use lossless codecs, especially as lossless JPEG 2000 has its issues in terms of resources and supporting applications. With FFV1 there are currently new options being explored and it can therefore be considered as not only forward-looking, but also viable for a number of color spaces and resolutions up to 24 bit (recent activities include discussion about FFV1 as a replacement for film DPX formats[7]). It presents a concept of storing mathematically lossless compressed data at a relatively small extra storage cost in comparison to the lossy compression alternatives such as IMX.

3.1.4. Partial file retrieval (PFR)

Retrieving footage from the archive and providing it for production is often a time-critical undertaking. News broadcasts in particular have to access content as fast as possible. Most of the time, short segments of content are ordered; however, retrieving the file stored in an LTO tape library may take some time. With high resolution or very long video files, the retrieval process can take many minutes, especially if there is a queue in the LTO library because of too few LTO drives. Even though PFR can be considered less and less important, with normalisations to one file format file sizes of up to several 100Gb are not unusual (e.g., normalised DPX sequences to FFV1, or analogue VHS legacy recordings in FFV1). This could be a single news clip that is part of an entire broadcast recording. Times that occur when retrieving this file from an LTO 7 tape are 11 seconds to load the tape in the drive, and a further average winding time of 40 seconds, but the crucial amount of time is the actual reading time with 12 minutes for a 104 GB FFV1 file (SD, YUV 4:2:2, 10bit).[8] All 104 GB have to be copied to the production storage and the content then has to be trimmed to the one minute requested duration. When AAM and Hierarchical Storage Management (HSM) allow partial file retrieval, just the requested segment will be retrieved from the LTO tape, reducing the waiting time drastically. Tape loading and winding times stay the same, but a 1-minute segment of the same content (about 0,85 GB) will be retrieved in seconds. How this performance enhancement impacts on production speed, especially in the context of HD and film content is obvious, when, for example, 50 requests per hour are the benchmark which needs to be fulfilled with typically 8–10 tape drives in the HSM system.[9]

7 Various public discussions, No Time To Wait! Symposium Vienna, Österreichisches Film Museum, November 2017.
8 "IBM LTO Ultrium 7 tape drive performance white paper," Support, IBM, accessed September 29, 2017, http://www-01.ibm.com/support/docview.wss?uid=tss1wp102594&aid=1.
9 "EBU Archives Report 2010", last accessed September 29, 2017, https://tech.ebu.ch/docs/techreports/tr006.pdf.

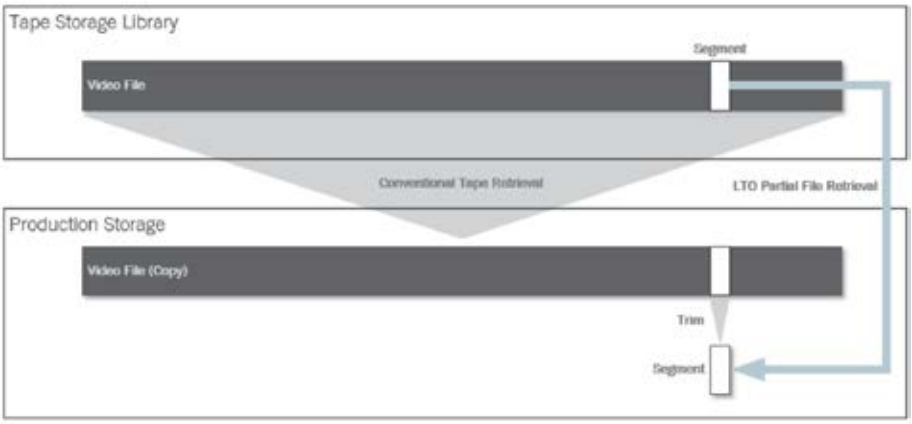


Figure 7. Schematic partial file retrieval.

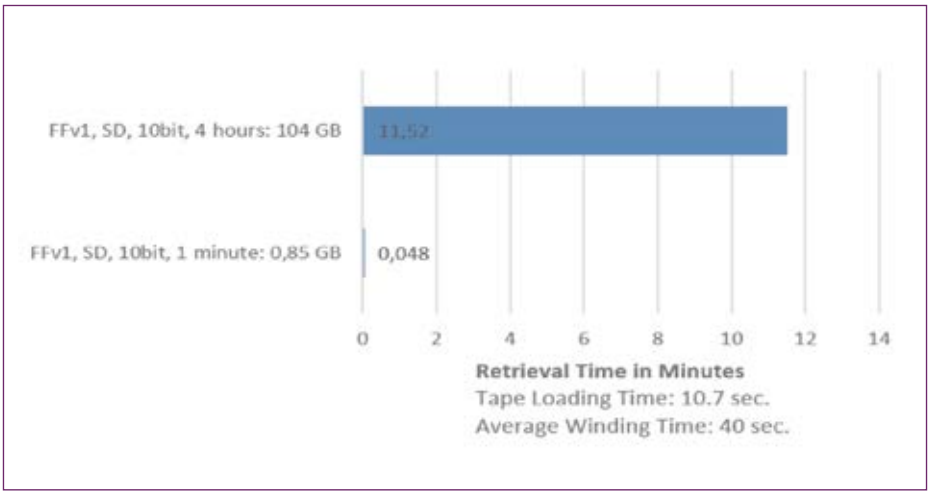


Figure 8. Comparison of retrieval times on LTO 7.

3.2 Metadata

In any library, the catalogue is the heart of the archive. Essence data retrieval in an archive is only as good as the cataloguing; meaning enriched descriptive metadata is required. With sufficient existing metadata, media becomes content. Adding rights information makes it an asset. The existence of metadata annotation makes the difference between an accessible archive with a well-structured catalogue and a black box full of unknown media.

3.2.1. Customizable metadata schemes

Broadcaster's archives typically have developed self-built metadata-only database schemes over time or may still rely on written cards as depicted above. Seldomly, has an archival standard found its way into that scheme. Our experience has shown that a variety of ontological understandings exist which may vary from carrier-based descriptions (typically) towards content-based annotations.

In any case, flexibility has to be applied (from archives, but also from vendors) to the fact that different archives can have different needs when migrating to an AAM. Broadcast archives in general will have quite similar necessities to each other, especially in comparison to libraries or other heritage institutions. An AAM has to be flexible enough to allow for the creation of extensive custom metadata schemes, reflecting the institution's cataloguing requirements.[10] As different item categories need different types of description, various logical entities (categories) will be conceived. Commonly used categories in the legacy archive sector of the broadcast industry are "Carrier", "Program", "Person", or "Title". Each category has to be defined at the field level with various options, including the name of a field, the type of data to be inserted (e.g., integer, character, or date) and the style in which the field can be edited (e.g., drop down menus, mask edits, or string edits).

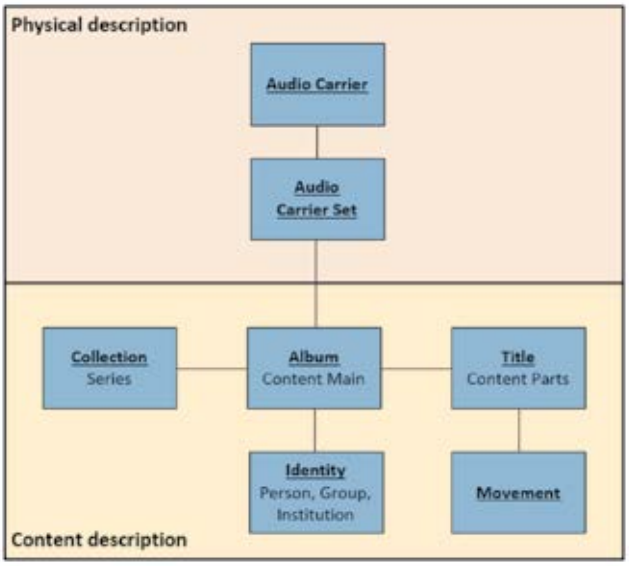


Figure 9. Example of a metadata scheme used for audio description.

3.2.2. Semantic metadata structure: FRBR and long tail search

To describe the ontological structure of objects and their relationships in a sufficient way, a strict hierarchical relationship between entities is not flexible enough, however by allowing an unlimited number of link types between entities that problem can be overcome. With qualified link types, people, content, documents, and contracts, for example, can be linked to each other creating perfect semantic metadata. It is a best match between object related descriptions and relational descriptions. Any object can be linked with another object by a qualified link. This functionality allows the implementation of many important metadata standards using FRBR (Functional Requirement of Bibliographic Records) such as the EN15907 recommended by FIAF.

10 Filip Kwiatek and Agnieszka Slomska, "How we managed our assets at NINA – The problems and challenges we have faced," In *Changing Sceneries – Changing Roles Part VII, Setting the standard in second generation MAM systems and metadata, 80-100*. Glasgow, FIAT/IFTA Media Management Commission 2015.

As a result, complex search queries can be defined for selectively finding content:

"Songs Edited in England with Mark Knopfler as composer, except between 1990 and 1991"
"Films directed by Fritz Lang, produced in the U.S. between 1936 and 1945, starring Spencer Tracy."

Defining such complex search queries is also possible by using Boolean operators, but doing so requires knowledge on the side of the user as a prerequisite. A user will just search for some (combined) attributes if he/she knows, or highly suspects the query will yield results. With an AAM that supports semantic linking, the possibility of serendipity is given. If a researcher can find content by lucky coincidence, without suspecting its availability, the hidden treasures of an archive can be unlocked as they are findable in a certain context. To a broadcaster this creates the opportunity to use unknown, rarely seen, or unexpected content for creative programming. By introducing semantic linking, the archive becomes increasingly mapped, and a user can practically navigate the archive using the established links while orientating him/herself on beacons of known contexts.

The image below exemplifies some possibilities of how content can be described when using related metadata. In this case a classical music album exists as an item in the archive. It consists of different parts that are related to different works. Various people act in different roles on the manifestation or work level. A user browsing the archive can enter the search at any known point and then explore the existing content without knowing about its existence, or about its inter-relations.

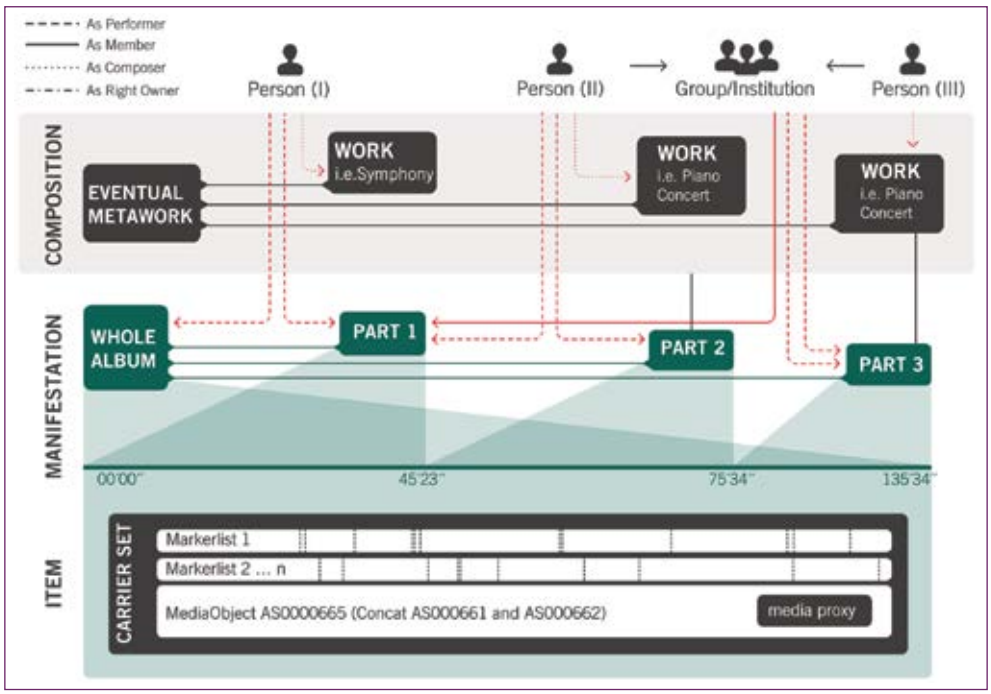


Figure 10. Metadata relations map the archive for semantic navigation – here as example from NOA mediARC.

3.2.3. Rights

One of the most difficult things when monetising content, is to have a cleared rights situation with archival content to make it a real asset. Whereas some countries do not have legal restrictions—or simply do not care—others struggle indefinitely with clearing rights for content, especially when exposing content to larger audiences. One of the functions of an AAM is to be able track and collect related existing contracts and to provide selective accessibility.

3.2.4. Post-Migration Consolidation and Revision History

When a repository gets migrated to an AAM, consolidation operations such as de-duplication have to be offered to solve possible metadata inconsistencies. During the lifecycle of a repository, metadata is likely to be changed over time. This implies the risk of losing correct information. Changes to metadata have to be automatically recorded as new revisions of the item. As a result, an item history is created over time. This way every former version of an item can be reviewed or even compared with another version. In this way, the cataloguing history of an item can be tracked and changed—metadata will not be lost completely if an error occurred during an inventory process.

3.2.5. Workflows

When coping with production systems and extensive ingest business processes with regard to both the file management and the workflow management, it is necessary that a workflow engine takes care of all information processing and workflow administration. Typically workflow management modules are located in 3rd party tools or are part of a monolithic MAM. Instead workflows could be allocated for all transactions that are OAIS relevant within the AAM, where OAIS consistency is also required. This guarantees consistency of the archive and all business procedures. This is especially important for the handling of large-scale automatic processes such as migration, mass ingest, delivery to production systems, or replication to cloud-based catalogues.

All interaction with essence data inside the archive has to be handled through preconfigured and monitored workflows. This is especially true for ingest, dissemination, and metadata enriching processes so that the archive system continues to conform to the OAIS reference model. Workflows of an AAM have to be designed to make business procedures manageable and plannable. Examples include delivering assets from an archive, validating metadata edits, or an extended QC process. A workflow administration area lists all running workflows and allows them to be paused, resumed, or reset and displays the current status of a workflow. During execution every workflow writes a log file for subsequent analysis.

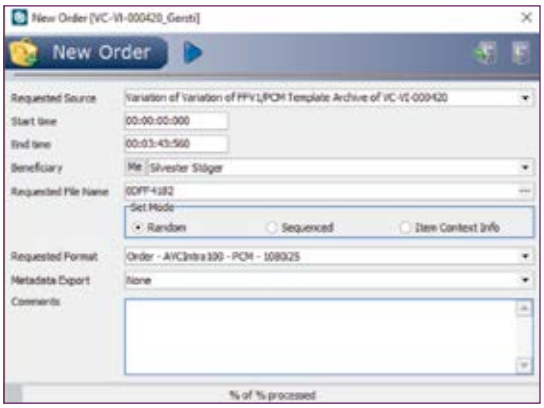


Figure 11. Workflow (NOA mediARC) describes visual interface to archive users to order from a lossless FFV1 archival format towards a desired target archive format.

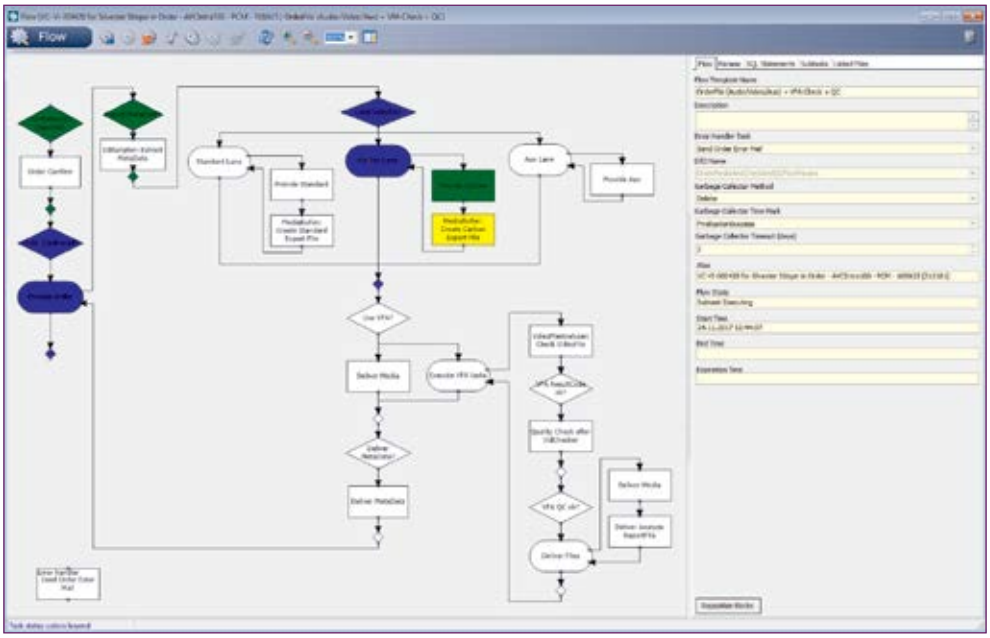


Figure 12. Workflow (NOA mediARC) describes the tasks in the background fulfilling export from a lossless FFV1 archival format to AVCIntra 100 incl. 3rd party video QC engine.

4. Conclusion

Broadcast archives need to preserve content and its stored value for the institution. In this context broadcast institutions become very similar to cultural heritage institutions that have to follow archival principles in order to safeguard their legacy collections. While one focus of broadcast institutions is naturally on their core business—production and acquisition of new content—the task of conserving self-produced unique archive content follows different principals.

A core topic for Archive Asset Management is the deployment of a specialized archive back-end following OAIS principles and focusing on the three main pillars: media, metadata, and workflows. These pillars include normalized file formats, coherent media management, semantic metadata description, and sufficient annotation that is key to good archive access in the future. Any AAM system that follows the OAIS principle will benefit from an integrated workflow engine available to provide consistent and automated business processes into and out of the digital archive repository.