

FAQ - CMS, DAM, MAM for Film Archives

1. What is the difference between a Computerised Database and CMS's, DAM's and a MAM's
2. Newer types of computer systems for organising, cataloguing and storing film archive collections
3. What is a Collection Management System (CMS)
4. What is a MAM?
5. What is a DAM?
6. Are there other dividing lines?
7. How do the needs for CMS for Film Archives differ from needs for CMS for non-audiovisual museum collections?
8. How are MAM's for Film Archives different from MAM's for television?
9. What are the criteria for selecting a *Collection Management System*?
10. Standards conformance: how to make sure that the CMS can handle the metadata model recommended by FIAF
11. Why is there no one ideal for all and a free Collection Management System developed by FIAF?
12. How much might such systems cost?
13. How to protect your collection/catalogue in the event of the resignation of a technology partner/developer?
14. Do I need the source code for the Collection Management System? Does having an open source Collection Management System solve all proprietary and financial problems?
15. Does FIAF recommend any specific product?

1. What is the difference between a Computerised Database and CMS's, DAM's and a MAM's

Within the practices of filmmaking, in the academic literature around the study of film, and in the marketing materials supporting the promotion of film we can find several broad categories and subcategories of systems which film archives can potentially use to organise their "assets" as they would say in the commercial sector, or "objects" and "collections" as we would say. Many of them are identified using acronyms: CMS (Collections Management

System), DAM (Digital Asset Management), MAM (Media Asset Management), ECM (Enterprise Content Management), to name a few. This can however lead to some confusion, as an acronym like CMS, which commonly denotes a Content Management System, is not to be confused with a Collections Management System. How do they differ from the databases that appeared in our institutions in the 1980s at the earliest, and more likely in the 1990s or 2000s.?

Computerised Databases, created in the third quarter of the 20th century, were derived from their paper-based ancestors. From the start early computerised systems shared many functionalities with modern CMS, MAM, and DAM solutions, primary among these being their ability to allow users to search and browse databases, or to easily allow the preparation of printable reports.

Over the course of their development these computerised databases allowed for networking with other locations and opening, for example, a text file. Such systems were created with analogue collections in mind, and not with “digital assets”.

What is a digital asset?

The two favourite words in the marketing materials for such systems are asset and content. Translated into the language of film archives, the asset is an object we can most readily refer to as a cinematographic work or an item. The asset being managed is stored in an analogue or digital format. In this latter case - a digital asset is detailed by its metadata. The metadata can detail the contents of the asset, the means of encoding, ownership, and rights of access to the asset.

The computerised databases are typically not set up to handle files (image and media), whether that be to upload them, or indeed to internally control digital files by attaching and assigning object metadata to a digital file. Furthermore such a database is not able to establish the continued existence of a given file, nor is it able to verify that what is described by the catalogue tallies with what was originally ingested.

Such a system could only link to a digital resource in the computer network, but did not “own” or indeed “contain” it. They did not allow the playback of video and audio material, as this had to be done externally.

Along with the development of technologies that led to the mass digitization of collections within heritage institutions, a need arose for systems enabling the management of digital collections to the same extent as analogue ones.

2. Newer types of computer systems for organising, cataloguing and storing film archive collections

Many contemporary CMS's offer institutions the possibility of centralising control and care over their collections, however diverse they may be, whether they be analog, digital surrogates, or indeed wholly digitally born items. In the beginning the functionality of most digital database systems could only readily extended to covering image based collections, hence why many museums were ahead of film archives in this regard. But today, along with the expansion of transmission and storage options, the functionality of modern databases can easily extend to covering the many needs of audiovisual collections.

Born out of the corporate world, MAM's and DAM's are capable of organising content stored inside of the system, but their focus is primarily on controlling ownership of assets (management and the logging of asset usage) and prioritising quick access.

To sum up all modern systems: CMS, DAM and MAM differ from previous more simple computerised catalogues in that they allow users to:

- Upload digital files / Import and export digital objects and their associated metadata
- Manage, organise, delete, and internally control digital files
- Attach data or object metadata to a digital file
- Quick search for objects and associated metadata
- Easily share objects and its metadata content

Additionally these databases should:

- Have preservation mechanisms in place
- Provide public access to the files through a public portal.

Typically DAM's and MAM's focus more on:

- Cooperation between departments and locations
- Facilitating the reuse of objects and its metadata / content

In corporate environments DAM is usually owned by the marketing team but can be accessed by multiple teams, including sales, product management, communications, or

other external partners. These different needs and approaches may potentially be "translated" into film archive workflows to meet the needs of accession, preservation, digitization, sales, and distribution teams in the institutional context.

3. What is a Collection Management System (CMS)¹

A Collection Management System (CMS) is a type of software used by heritage institutions, primarily museums and archives, secondarily by libraries and galleries. Collection Management Systems allow users to organise, control, and manage objects within their collections by "tracking all information related to and about" those objects.²

The purpose of a CMS is to take care of an institution's entire collections, both analog and digital. They connect recorded cataloguing information to the objects either stored outside the system (via signature number/id number and location information) or link to its digital version within the system.

"CMS grew out of catalogues, more recent and advanced systems are being used to improve communication between museum staff and to automate and manage collections-based tasks and workflows. In larger institutions, the CMS may be used by collections staff such as registrars, collections managers, and curators to record information such as object locations, provenance, curatorial information, conservation reports, professional appraisals, and exhibition histories. All of this recorded information is then also accessed and used by other institutional departments such as "education, membership, accounting, and administration."³

Collection Management Systems are also used to provide access to information about an institution's collections and objects to academic researchers, institutional volunteers, and the public, increasingly through online channels.⁴

Collections Management Systems can now incorporate DAM or MAM functionalities by tracking and completing the workflow of digital preservation tasks. Such systems will support:

- Digital file storage

¹ Throughout this document we use the abbreviation "CMS" for "Collection Management System". Do not confuse it with "Content Management System", which, outside archives, museums and libraries, is more commonly associated with the acronym "CMS". Content Management Systems are an entirely different matter and should not be confused with Collection Management Systems. .

² https://en.wikipedia.org/wiki/Collections_management_system#cite_note-3

³ *ibid*

⁴ *ibid*

- Digital file preservation
- The ability to assign data to a digital file
- The ability to contribute object metadata and link it to a digital file
- The ability to provide public access to both object data and content.

4. What is a MAM?

Media asset management system (MAM) is a single repository for storing and managing video, audio and multimedia files and their associated metadata.

Media asset management originally emerged as a video/audio storage and management solution for the television and movie industry to manage large-format video and audio files and production and distribution workflows. Since it was originally created with TV in mind, the emphasis will also be on providing permission-based access to content and to facilitate precise quality control for the purposes of clearing material for broadcast.

MAM comes with a separate transcoding tool, also called “video conversion software” or a compressor. The function of the transcoder is to change the format of the codec (convert) and container (re-wrap), it can change aspect ratios, change layout (slate, bars, etc), it can extract a fragment, and it can insert logos or text. A transcoding tool must be chosen in consultation with the system vendor to ensure complete integration. The MAM supplier will probably already have transcoders that easily integrate with their systems. As with any other software transcoders can be open source or proprietary.⁵

While a MAM provides in-depth management of video, this solution was originally not intended to manage, preview and distribute smaller media, like documents or images. Film archives not only hold audiovisual collections but also a mass of various non audiovisual film-related documents and images.

Consequently this demands that institutions pay close attention to whether a MAM can meet all their needs for digital cinema and video management, as well as other functions needed in the wider scope of an archive’s collection management processes.

⁵ You have to take into account the potential patent rights to file formats in some countries. For example, the EU and the US have different laws and views on copyrighting the algorithms.

5. What is a DAM?

Digital Asset Management can broadly be defined as a system for organising, storing and retrieving rich media and managing digital rights and permissions.

Traditionally DAM systems could be viewed as systems primarily centred on brand assets, such as logos, images and other media files, and their reuse within an organisation, much like a library asset management system.

Originally these systems were an integral part of the media and content creation process, helping marketing teams organise, collaborate and distribute their visual media from within one central library of digital assets. As such they focus on organising marketing materials including document files (DOC, PPT, XLS, PDF), product photos (JPG, PNG, TIFF, RAW, GIF, BMP), and more recently 3-D files. Although modern DAMs support the most common video formats (MOV, MP4, MKV, AVI), and audio files (WAV, AIFF, PCM, MP3), it is very unlikely they will be able to support less popular professional formats.

Depending on the system and its additional integrated tools, users can perform a variety of functions, including re-sizing & re-formatting images and video. An advantage is that they are often “format agnostic”, meaning that users can upload nearly any type of digital file. Whether you can view or play this media correctly is another question. Whatever the final destination is, the goals of DAM are the same: find content quickly and put it to use. A DAM can provide tools for: tracking usage history, sharing large files easily with external parties, version control and edit tracking of all media assets, access and permissions management for internal and external users, compliance regulation, as well as advanced recovery and geo-replication. The permanence of the asset/object and its description is neither the focus nor the function of a DAM.

Off the shelf DAM systems are not readily designed to freely create metadata structures with the media and content it incorporates.

6. Are there other dividing lines?

The next fundamental choice is an archive’s **approach to storing any and all digital collections**. This is divided between the more traditional approach of using locally-hosted solutions and the more recent option to rely on cloud-based asset management systems.

- 1) The first requires building a digital repository (hardware and software) in your archive. You have to buy everything almost at once and make a lot of strategic

decisions from the start. Beyond this initial hurdle, as owner you have more control over both the technology and all future expenses, which decrease significantly after the initial outlay.

- 2) The second allows users to access assets from almost any device and location. Cloud software provides an “abstraction” to the storage system that you are using, but you are not the owner of either the hardware or the software. Using a cloud-based service means less time spent on managing the system, instead outsourcing this problem. It can mean no control or knowledge of either the data storage or processing locations, as you will probably not be assigned any specific hardware solutions.

Both solutions have severe financial consequences, very different cost curves, and have different implications for financial, legal safety, and disaster risk management.⁶ Especially important to remember is that you and the cloud system provider have clearly different interests and goals.⁷ What is initially attractive with cloud storage systems is that they require a relatively low financial outlay to set-up, and that payments are fixed and grow in proportion to the growing collection. However, the long-term cost of such a solution for large installations will most likely be much higher than building your own archive in-house.

The next dividing line will be your **approach to accessibility and network architecture**. The choice is primarily between having the system installed directly on your computer and only partially on a server or having it all based on a central server.

- 1) In the first case, originally known as "fat" or "thick client" can also operate partly “offline” (you are connected to a server over a network but don't need to have a maintained connection). In such cases computing power is split between server and client.⁸
- 2) In the second case, also called "thin client", popularised in the first decade of the 21st century, is a system in which a user accesses a system through a computer, depending on a server's applications and its computing power and you don't need to

⁶ In the event of a sudden decline or loss of financing for your catalogue and digital collections, you can even passively store data for some time. In the case of cloud solutions, this involves the risk of data being deleted by the host/ hosting service provider.

⁷ In an extreme case, it may be cheaper and more attractive for commercial cloud storage/hosting providers to pay a penalty for losing your data, rather than provide stricter data security, which would mean higher costs for them.

⁸ This solution has clearly changed over time - in the beginning of computer networks on the server side there was only a database, later consequently, more application logic went from the client side to the server side. The consequence of this tendency was the evolution into today's web browser-based applications.

install any additional software and it is OS independent because you need only a supported browser. Attention must however be given to noting supported browser applications and versions though, and ensuring that updates are provided to adapt to new browser versions in the future.

7. How do the needs for CMS for Film Archives differ from needs for CMS for non-audiovisual museum collections?

Since nowadays the possibilities and functionalities of individual classes of products are blurring, one needs to pay attention to how a given system deals with audiovisual media, regardless of the software class (type) it represents.

The vendor or supplier of any given system will readily tell you that they all "support multimedia", but the devil is in the details. CMS software typically used by (non-film) museums specialises in describing and managing 3D objects or digitised 2D objects. These products are often not tailored to film and video collections and the broad spectrum of possible digital audiovisual formats.

The acknowledging the difference for software providers between non-audiovisual museum collections and film archives collection of various formats digitised film is essential.

Audiovisual files are bigger and the content is time based, which means it needs a powerful network, considerable computing power, and a flexible transcoding tool (Video Conversion Software) integrated into the system, and a smart video player to access material in the system. A versatile video player is essential for film archives, as cataloguers have to deal with irregular frame rates and they need accuracy down to the single frame, the ability to play video backwards is vital, and not as easy task as one might think.

What makes digitised films so complicated to handle is not only that the files might be extremely large but also that they can often be in complex associative relationships with multiple other files and films. Film archives pack their digitised film elements in wrappers with attached image, audio/sound, and subtitle files. In film archives you will also find file sequences with hundreds of thousands of files (DPX, TIFF, JPEG2000) which have been used to create both preservation and master copies.

Likewise these systems tend to have problems handling lots of relatively small files, (e.g. DPX) because this need was not originally considered by the system's designers.⁹ It is typical for film archives to operate on a very high numbers of files that can contain only a single frame each.¹⁰ Queuing such tasks for several movies, saved in image sequence format, will be an extremely good test of system performance and design.

In any case uploading and watching high resolution files clogs bandwidth, are difficult to handle, and extremely slow to move from one location to another. Also you cannot readily watch your TIFF sequence from your CMS on your laptop. That's why these systems regularly use smaller representations of these files, proxy files, for viewing. You can generate them automatically while importing, or create them externally. Moreover, how your system handles video files sets the bar overall, as these formats have a lot of technical metadata to display and keep unchanged.

In summary, CMS and DAM provide functionality for managing a wide breadth of digital assets, including video, but it isn't tailored to the needs of the broadcasting industry in the way a MAM is, including multiple valuable functionalities for film archives. While a DAM can provide some management and workflow options for video, a MAM is designed specifically for delivering workflows in high-volume video production. Pay attention to whether the system you are interested in allows the correct interpretation and reading of formats used by your film archive, codecs, wrappers, or can display untypical frame speed and aspect ratio. Check if technical metadata of preservation (and other) files are not changed after ingesting, and displayed correctly. Do you have the option of transcoding media for production and access (proxy) and does the system offer reliable storage for preservation files? Bear in mind the existence of open source tools which you can add to your Collection Management System, such as MediaInfo and the FFMPEG family of tools.

8. How are MAM's for Film Archives different from MAM's for television?

Television archives and film archives do store similar digital audiovisual formats, but outside of this part of our collection further similarities are much fewer. Broadcasters use a lot of tools that film archives don't necessarily need. Therefore be careful not to pay for expensive solutions that you will not use at all.

⁹ This is true of all off the shelf systems

¹⁰ e.g. the process of copying 1 million files one by one can instigate a cascade of checksums to see if they still match the originally ingested files.

Cataloguing standards

As film archives we also need to think also about cataloguing and organising non-film collections and the best way of expressing their relations to cinematographic works. Both types (television, and film archives) are focused on accurate metadata and cataloguing information but we have different cataloguing standards, traditions and focus points.

For MAM developers more accustomed to working with broadcasters, FIAF's cataloguing standards present a new unknown, and the method of expressing CW in 3-4-hierarchical levels can appear overly complicated, and one they cannot readily accommodate. It is very likely that a MAM system does not allow for deeper editing of the metadata structure. (More in [no 10](#))

Supported formats

Film archives are not focused on reusing assets and their focus is on protecting data from change rather than facilitating and tracking it. One of our key tasks as film archives is digital preservation, supporting both lossless formats and formats specific to film production. As part of our legal deposit obligations, we receive a lot of films that have been born digitally. We are obliged to protect the technological origin of the film, the original characteristics of the film and appendant/associated technical metadata. This is another difference between film and broadcast archives, where the latter normalise storage formats and codecs to simplify tasks and production workflow.

For preservation reasons in film archives some formats, which are rare in broadcasting production, are in use (uncompressed formats, file sequences like DPX, higher bit depths, logarithmic files). They likely are unsupported by off the shelf TV MAMs, as they are definitely not required by broadcast clients. And even if they are supported by the integrated transcoder (MAM does not have to display files directly), CMS/MAM/DAM may have a problem with how best to display the file sequence for the movie in the user interface (for one 3h movie, it will be about 250K files).¹¹ From this perspective film archives are a more complicated client for software developers and you must make sure that your key needs will be fulfilled.

Emphasis on shared access editing and production

¹¹ For example, a typical pre-built system might want to display them all in gallery view with the entire metadata list next to it, and in the process crash the whole system..

A MAM system has traditionally been the centrepiece in file-based (video or audio) workflows for broadcasters. It is part of the A/V production chain, and they also often have the most comprehensive integrations into the tools that production teams are already using. For example, allowing editors to store their content in a central location for shared access by many co-workers using non-linear editing tools, transcoding tools (video conversion software) and automated quality control (QC) tools. Here we have the next major difference - the desire of broadcaster's is for a workflow that is immediate, and which offers and shares simultaneous access to the same file in order to facilitate collaboration in editing. Whereas the archive wants to keep files in their original form, at most facilitating any editing work well outside their protected system. Rather, the system is to be a guarantee of the authenticity and immutability of the file as an archival object.

Integrations with third parties, compliance to a specification

But remember that some of these tools designed for broadcasters may be useful and may only require adaptation. For example, our QC protocols are not designed to meet broadcast's need for "compliance to a specification"¹² or their need to identify problems before delivery or going to air¹³. What the archives have in their collections is often in unusual formats and to varying degrees of degradation and technological advancement - so television QC would immediately deem them unsuitable, and that is undesirable. In archives it is still necessary to involve a human operator reviewing the file as a form of "manual QC." But other forms of QC require somewhat less technological involvement (i.e. correct file type, gap in numbering files in sequence, file size differences, file errors) and more advanced tools can demonstrate their utility for detecting artifacts in digital images, video, as well as audio.

9. What are the criteria for selecting a *Collection Management System*?

You can start with the established criteria used by Ashley Blewer in her live [Collection Management System](#) spreadsheet list. [Here](#) is more information about the list.¹⁴ What else should I be careful about when choosing a CMS product?

¹² Compression and formatting for the specific requirements of the TV signal or VOD

¹³ Correct time code, correct layout (slate, bars, etc), colour gamut, audio errors (loudness, clicks, silence)

¹⁴ <https://bits.ashleyblewer.com/blog/2017/08/09/collection-management-system-collection/>

Project success and its durability/stability depends upon the understanding of what we need, allocation of adequate resources, time and budget.

Below are considerations articulated by The Canadian Heritage Information Network, meant for museums, but these could also be used universally by film archives:

Know your needs

- Make sure needs have been closely examined and documented
- Do not look at systems without knowing what is needed. It is too easy to get caught up in what is demonstrated and not realise the system isn't really what is required.

Vendor Considerations

- Are they reliable?
- Have they installed other systems?
- Have they installed similar systems?
- How long have they been in business?
- How accessible are they?
- Are they easy to talk to?
- Are questions answered clearly?

References

- Check references thoroughly.
- Talk to others that have had the system installed.
- Try to visit a site and see the system in operation.

Support

- What kind of support does the vendor offer?
- What kind of documentation is available?
- Are others in your area using the system?
- Does the vendor install the system?

(...)

Technology Issues

- Does the system run on specific machines?(...)
- Is the technology up to date or is it at the tail end of a technology phase?
- Updates to keep up with technology

- Is the technology compatible with the rest of the museum in order for information to be exchanged in the most efficient manner?¹⁵

Additional technological considerations for Film Archives:

- How much user interaction is managed by open Web services.
- Does the system allow for communication with existing/future partners adding or modifying input/export forms - API and its documentation (again).
- Can you create a test environment to perform tests on it, in particular migrations and updates, train new users? Ideally before you purchase a product to formulate the decision.

Workflow Management tools. It would be great if the software(s) you use could represent and support the workflow of your organisation and offer adequate views for different groups of users.

- Can essential elements of your workflow be covered and integrated by this/these tool(s) (you could think about e.g. accession, conservation, cataloguing, documentation, digitisation, distribution, online publishing, and digital asset management and preservation).
- Flexibility on creating and customising report formats (formatted text, PDF & spreadsheet reports). Will the system be able to assist in the transfer of information inside and outside of the organisation (accession reports, preservation/ restoration reports, film material distribution and loan reports).
- Tools allowing you to manage the flow of items and location tracking.
- Registering the editing history of metadata.
- Data export to various standard formats.

Background media workflows and “batch services” may be important for those archives that have a large file collection and a wide flow/stream of digital materials. Such functionalities will allow the delegated administrator to process many files on the central server. These operations take place in the background and do not burden the individual workstations. Batch transcoding, cataloguing, updating, editing and importing features allow the archivist/digital media administrators to queue and make automatic changes.

¹⁵ Shona Fitzgerald, The Canadian Heritage Information Network's: *Collections Management Software Review*, May 27, 2016

Web Publishing tools for media - if you need it, check what you are getting.

Handling of digital film formats - look at questions [from 2 to 5](#).

Flexibility in Data Modeling. Customization and perfecting metadata schema is a never ending process. The needs of describing your collection will be evolving and you will change schemas, you will add fields and define new values in vocabularies. Consider how much you can change on your own and for which you have to pay for externally. The work of a developer on the part of the supplier will cost you more, and cannot always be done immediately and you have to spend time explaining what you want. Additionally, the product should support the recommendations of the FIAF Cataloguing Manual - For more on this, see question [number 10](#).

Security, access and granular control

- You will probably need user access groups and permissions to protect your metadata from unapproved change and copyrighted material from unauthorised use.
- Establishing and changing access rights should be easy to do in the organisation (you don't want to pay every time you change the rules for a particular group or user).

Search tools. You may be interested in the following functionalities:

- Full text search and indexing.
- Catalogue and cross-reference everything, from object data to constituents.
- Configurable navigation tools, especially query filtering and advanced search operations with more freedom in building and displaying more complicated queries.
- Pay attention to whether your language is properly supported. Check if values with special characters are searchable and sorted in the correct alphabetical order.

The beginning and end of work in the (new) system - See question [number 13](#).

Costs - look at question [number 12](#).

10. Standards conformance: how to make sure that the CMS can handle the metadata model recommended by FIAF

Can your chosen CMS implement 2-3-4 levels of a description hierarchy? Only then can you successfully fulfil the requirements of the FIAF [Cataloguing Manual](#) and Standards for the description of cinematographic works - [EN 15907](#).

There is a risk that suppliers will promise you that without seriously engaging with these standards first. In reality what looks easy, can prove less so, and you should consider checking if:

- Is it possible to build hierarchical, as well as non-hierarchical, relationships/links between fields and objects?
- How does data structure influence the result and the display of search results?
- Can fields be repeatable and nested?
- Can field values be displayed in lower and higher levels of hierarchy (virtual fields)?

If you want to make sure that a software solution supports complex hierarchies for objects, there is no other way than building such a schema. Consider trying to enter a 3-4 level object which expresses possible relationships between fields and values (using FIAF Cataloguing Manual and Standard, for the description of cinematographic works - EN 15907 as a model). Check that the data is displayed correctly and try the search tools provided.

11. Why is there no one ideal for all and a free Collection Management System developed by FIAF?

This question comes back to members of the CDC on a regular basis. The FIAF community and its committees are not corporate bodies with the competence/knowledge, financing and *modus operandi* that could support such an endeavour. It is also not our function.

Developing a software product is a large, long-term business investment and must be done with the same resources as commercial enterprises provide.

It is true that heritage institution communities have been given the framework for such an undertaking in the past, but it happened in the field of basic computerised cataloguing software (see question 1). Developing a CMS (or MAM/DAM) with built-in collections

management functionalities and with advanced digital preservation and media management functions is another story. The second requires a great deal of time and even more specialised expertise, both in the design and development phase and in the longer term production and maintenance environment. FIAF does not have money nor the human resources to answer this challenge responsibly.

Moreover there would be no one client to negotiate how such software should be working, how the GUI should look etc. It would not give a chance to select the best solutions and eliminate the less successful ones through open competition.

By sharing knowledge about best practices and, and by networking between FIAF users and the many different existing solutions, we would like to help you learn from our members good and bad experiences with already existing software products.

12. How much might such systems cost?

In addition to understanding your own needs, and surveying what vendors/producers offer, there are three more things to be taken into account:

1. Money (as cash that must be spent at the beginning of implementation). The choice of adoption of open source software has a lower up-front cost and shifts the cost centre from licensing to customization and implementation.

2. Money (you spend on maintenance every year during its use). A system cannot be developed and then simply be left to a life of its own. A maintenance plan is essential to provide for support, ongoing training, and to ensure file maintenance, sufficient backups, changes, enhancements, etc.

It has been assumed that for proprietary solutions the annual support is between 7% and even 20% of the cost of the software purchase, and in the case of open source solutions the cost depends on the model and how active the associated community is, and your organisation is to enhance and update the product continually.

3. Money (for staff time). You will need (in-house) hardware and software administrators and metadata cataloguing managers. Probably you can reduce the cost of point 2 by developing the scope of competences of admins from film archive to handle such matters and in turn limiting work left to the vendor's own developers.

Watch out for the small print! The system may also be more expensive than you first think following initial conversations with any given vendor, because of various and not always transparent licensing models. One way of licensing is by installation - for an unlimited

number of workstations and user accounts. Second is based on the number of concurrent user licences. Therefore, find out if the use of the product you are considering is scalable. System could have the cost of licensing proportional to the amount of data stored in the digital archive. Such models you can find (but not only) in the field cloud storage solutions. Prices can be calculated annually, per month or as a perpetual (lifetime) licence. It can be priced per 1 user or with unlimited users. It can also rely on a combination of criteria and different units (1 user per month; 50 Users Per Year; etc.).

Take your current and future needs to calculate the cost curve.

What else can we think about? *Canadian Collection Management Review* proposes following list of questions to be considered at the start:

- Is data entry included in the price?
- Are there charges for changes to the system?
- What training is included in the purchase price?
- How much do upgrades to the system cost?
- Are there charges for trouble-shooting or start-up?
- Can there be a holdback in payment until the system is accepted?
- How much are additional copies of the system¹⁶ or Sandbox.¹⁷

13. How to protect your collection/catalogue in the event of the resignation of a technology partner/developer?

It can never be ruled out that our technological partner will resign at some point, or that we will no longer get along with him or her. In a worst-case scenario, software products may become suddenly abandoned by the primary developer. One method of minimising this risk is to use open source elements. Many institutions have systems which avoid proprietary software but they still have some proprietary components such as device drivers (think of film scanners, colorimetric devices, etc), it can also affect communication protocols, messaging, or data formats. Whatever your CMC choice was, you might need to know how you can expose application logic without the help of the original software developers, and in turn access to source code.

¹⁶ Collection management software review p. 9, 11

¹⁷ After installing a system it is advisable to have a separate testing environment (often called a 'Sandbox') to perform metadata schema changes, test migrations of data to the system and upgrades.

Even when the cooperation between vendor and institution is excellent, we should always prepare for a rainy day, by:

- product documentation supplied by the developer
- documentation of customization and scripts in use
- make regular backups of your database
- access to the system via API as well as good API documentation.
- be able to export the entire database (and have a test export performed). You should also be able to extract the entire database in a human readable format (like XML or JSON). Nevertheless in the common case of relational databases the best solution will be an SQL dump. On such a basis, you can restore the entire database and generate whatever you want from it (as well as the aforementioned XML or JSON formats).¹⁸

This way, you should be able to generate all the metadata and relationships between them, subject specific dictionaries and files related to the metadata. End-of-contract data extraction can also be entered into the contract with the software owner. But for the sake of security, it is worth knowing how to do it yourself.

14. Do I need the source code for the Collection Management System? Does having an open source Collection Management System solve all proprietary and financial problems?

There are many reasons why safety and access to entrusted materials, and the artefacts of our collective memory, should be based on open and transparent software. This allows institutions to control expenses, protects them from the potential consequences of the resignation of any primary developer and, above all, gives you the freedom to study, modify, innovate, and distribute what you use.

Also, the EU and other governments have been advocating the use of open source solutions for two decades.¹⁹ You can also read about open source solutions in the realm of transcoding in FIAF sources,²⁰ or find many individuals and non governmental/non-public bodies advocating for openness of code.²¹

¹⁸ Reconstructing an entire database from aggregated views in XML or JSON may not always be feasible.

¹⁹ https://ec.europa.eu/info/departments/informatics/open-source-software-strategy_en#opensourcesoftwarestrategy

²⁰ Reto Kromer article in JFP 96/2017: https://www.fiafnet.org/images/tinyUpload/Publications/Journal-Of-Film-Preservation/Matroska-and-FFV1_Kromer_JFP96.pdf.

²¹ <https://opensource.com/resources/what-open-source;>

We should nevertheless remember that open source doesn't come without a cost, even if it's "free" and the relation between proprietary and non-proprietary systems is not a black and white issue. Open source software producers are often private companies that earn on additional software elements like enterprise-grade features under proprietary licences, or other services created around the use of free products.

There are a few issues you can think about when your strategy is to use open source and/or your goal is to have source code in your Cataloguing software CMS/MAM/DAM:

- Is there an open source product(s) satisfying your needs?
- Does it draw together an engaged community of developers with similar goals to yours?

If not, and you want to be the owner/holder of the source code, and/or have a software based on open source, you can also:

- A. purchase and develop a customised system for you (preferably also on existing open source elements), together with full or limited rights to the source code for it.
- B. develop a customised system in-house, using your own developers.

As you are a film archive, software development probably isn't your core competency.

Therefore the further questions might be:

- Can you be technologically better (in particular hire and keep good application developers) and economically more stable than a company whose business is fully focused and devoted to creating software products.
- Are you ready to bear the risks that in a commercial partnership would be on the side of the vendor (i.e. to complete the tasks within the assumed time and budget)? Remember that developer costs are high and constantly growing.
- Whether you will be able to develop and support creators for your product independently from/of the primary developer or without a community of open source developers (when your software grows and changes separately, creating a distinct version, a so-called "fork", and financing its independent development).

Whatever your choice is and even if you have decided to use proprietary software in your CMS/MAM, there are still a few things to be avoided:

<https://opensource.com/life/15/12/why-open-source>

- Proprietary technologies and application-specific components, especially in the area of databases.
- Proprietary non-application-specific components like file formats or deep archive recording formats, e.g. recording formats on popular LTO tapes should be open source.
- Using software that hardly anyone uses, just like an open source product you also need a community/partner(s) to discuss development.

What is recommended/preferred:

- software solution(s) which are modularly constructed, improving both the flexibility of further development, and the robustness of the code. This lets you add new modules and exchange parts and include open source software components;
- modules used in your software solution(s) preferably should be based on existing up-to-date, popular and state-of-the-art open source code elements (or closed but well documented). It will guarantee your solutions a longer life;
- nonproprietary communication protocols, messaging or data formats;
- negotiate with the developer/vendor a safe way out and include end-of-contract resolution in the contract;
- transfer more control over (metadata schema and fields formats/values and general maintenance) the system, from the vendors developers to your own system administrators;
- make good documentation, in case you have to opt out of the application used by the institution - see more in the answer to question [no 13](#).

15. Does FIAF recommend any specific product?

Every film archive has its own profile and collection needs, and out of each institution's individual technological context the overall needs, as well as access to resources and finances will be very different.

Collection Management Software can take a different place and shape, depending on your collection and whether you are involved in a digitization project and constantly producing new digital entities/items, in which case, you may need to pay more attention to digital preservation tasks. Needs also depend on existing surrounding hardware and software architecture.

For these reasons, we see our task rather in:

- reducing any duplication of effort (and costs) across the FIAF community by providing a platform where archives can share information and gain expertise;
- promoting the innovative and collaborative power of open source solutions;
- creating guidelines that will help in navigating between possible technical and business solutions;
- promoting the importance of standards regarding cataloguing and digital preservation;
- helping you implement good practice, in whichever software and hardware solutions you have chosen;
- helping you evaluate long-term maintainability of your solutions and its standards conformance.

Written by Elżbieta Wysocka, 9.2022

The author would like to thank Detlev Balzer and Georg Eckes for the reviewing draft of this article and their helpful comments, as well as Adelheid Heftberger and the FIAF Cataloguing and Documentation Commission members for their kind support.