Digital asset management and libraries, archives and museums: Separation and convergence

Helen Hockx-Yu

is a Program Manager at the Office of Information Technologies, University of Notre Dame, where she coordinates a campus-wide strategy to guide the preservation, access and dissemination of digital assets. Her previous roles include Director of Global Web Services at the Internet Archive, Head of Web Archiving at the British Library and Programme Manager at the UK Joint Information Systems Committee. She has led many national and international digital initiatives, and has written and spoken extensively about web archiving and digital preservation, addressing national and international audiences at various academic and professional conferences.

Office of Information Technologies, University of Notre Dame, Notre Dame, IN 46556-4655, USA Tel: +1 574 631 0533; E-mail: hyu3@nd.edu

Abstract The University of Notre Dame’s latest Campus Crossroads Project includes the Rex and Alice A. Martin Media Center, with a 186 m² studio, which has brought together previously dispersed media production work across campus. This is also an opportunity to develop a common solution for the life-cycle management of video assets, including archiving workflows and digital preservation requirements. This paper discusses the differences and similarities between the digital asset management (DAM) and libraries, archives and museums (LAM) communities, and the software used by them, in the context of a project aimed at assessing requirements and making recommendations for video assets management at Notre Dame. It argues that DAM and LAM professionals have a lot in common, and that the modularisation and convergence of DAM and digital preservation systems offer the best of both worlds and life-cycle support for digital assets in diverse environments such as higher education institutions.

KEYWORDS: digital asset management (DAM), media asset management, archiving, digital preservation, libraries, archives and museums (LAM), file fixity, file formats, metadata, software modularisation

INTRODUCTION

The University of Notre Dame (UND) is a private research university located adjacent to South Bend, IN, USA. It was founded in 1842 by Father Edward Sorin, a priest of the Congregation of Holy Cross. UND was ranked 15th among national universities in 2017¹ and 12th among the top research universities in 2016.² Catholicism and Catholic intellectual tradition continue to serve as the underlying foundation for the university’s aspirations as a community of scholars and administrators. It is also the university’s point of greatest distinction from many other research institutions.³

Digital media are frequently produced and widely used at UND for education, research, special events, faith-based services
and athletic competitions and performances. UND’s media products range from photographs and simple sound or video capture to sophisticated footage appropriate for national broadcasts.

UND’s media production services are traditionally provided by multiple groups on campus, each with its own business model, funding approach and management and personnel. While this builds specialised expertise and meets the requirements of specific types of users, there are also disadvantages. Video assets, for example, are stored in different locations and managed independently, making it very difficult to discover knowledge, locate and reuse the terabytes of media content produced on campus.

UND is currently in the last phase of the Campus Crossroads Project, the largest building initiative in the history of the University. The project is aimed at maximising the potential of one of the most recognisable and centrally located buildings on campus: the Notre Dame Stadium, which is now also the home of the Rex and Alice A. Martin Media Center. It includes a 186 m² studio, bringing together previously dispersed media production work across campus.

The establishment of Notre Dame Studios (ND Studios) provided a strong impetus for developing a common solution for the management of video assets, in the same, coordinated manner.

THE VIDEO ASSET MANAGEMENT ASSESSMENT PROJECT
Jointly led by the IT department and the University Libraries, a project was initiated in Spring 2017 to make recommendations for video assets management. The project involved departments across campus that produce videos or provide technical services or custodianship over videos, with the goal to reflect upon video use-cases and articulate requirements. This is a step towards the implementation of an enterprise-wide solution for managing Notre Dame’s video assets, which are estimated to measure c.2 PB at present, and growing. The solution needs to meet a number of essential requirements:

- **preservation**: the security of knowing institutional heritage will be available for generations to come;
- **accessibility**: the ability to easily search, find and use elements of videos, if authorised, inside or outside the university;
- **financial stewardship**: the deployment of the most economical solution possible; and
- **shared expertise and streamlined workflow**: the aggregation of human talent and technical infrastructure to benefit the entire institution.

A digital asset management (DAM) or media asset management software was an expected component of the solution. The solution also goes beyond technology and includes policies and business models that link components of an ecosystem of solution.

THE STAKEHOLDERS
The project worked with campus stakeholders representing multiple departments. They can be divided into the following broad categories, each having distinct workflows and requirements:

- **Content creators**: Videos produced at UND cover a wide range of academic, athletic and faith-related content. Producers, videographers and editors require tools for capturing, logging and editing videos, as well as for the client review and feedback process. They require search capability and ready access to work in progress and finished content. They also need to access inactive and historical content, so that this can be repurposed or reused.
- **Technical service providers**: As the campus technological provider, the IT department is responsible for the operation and
maintenance of most technical services relevant to video assets, including live streaming of lectures, and various campus and sporting events. IT also pays a great deal of attention to the non-functional aspects, and requires that information systems are secure, maintainable, reliable, usable and meet certain performance criteria.

- **Custodians:** Librarians and archivists are custodians of the university’s videos assets, whose long-standing role is extended from the analogue world. The oldest known moving images of Notre Dame, for example, preserved by the University Archives, date back to 1912. To ensure videos of enduring value remain accessible and usable, librarians and archivists require provenance and authenticity assurance, file fixity, risk identification, preservation planning and content migration support. All activities that have had significant impact on the trustworthiness of video assets need to be tracked, recorded and can be reported on.

- **End users:** End users, who could be external to or affiliated with UND, consume videos made available by UND by searching, browsing, viewing and downloading them, with or without payment. End users are typically not involved in altering the intellectual or technical characteristics of video assets. Videos are published and delivered in many locations, streamed live or on demand, including learning management systems, social networks, content networks and hosting platforms.

- **Researchers:** Videos are frequently used in social and medical research at UND (eg for behaviour observation). These are generally under the control of faculty conducting the research and have limited needs for sharing or archiving beyond the projects. The requirements of academia, however, in addition to general storage and information management needs, are diverse and complex, and differ across disciplinary areas. The project did not focus on supporting research use at this stage.

**THE VIDEO ASSETS LIFE CYCLE**

When pulling together all stakeholders’ workflows, a full life-cycle picture of video assets emerged (Figure 1). This provides a useful framework that allows the identification of essential processes in the lifespan of video assets from creation to preservation. It also recognises the change of focus and frequency of use over time, and can be mapped easily to the group of stakeholders performing each process and/or interacting with video assets.

As shown in Figure 1, ‘capture & production’ and ‘post production’ are the essential processes related to the creation of video assets. More than simply the action of shooting and editing videos, these processes also include idea generation, planning and development, and client feedback — all those things in the creative process that transform the idea to the finished product.

‘Publish & deliver’ is the process where videos are disseminated to various locations or destinations to serve end users.

‘Access & use’ refers to the process of end users consuming videos.

‘Archiving’ refers to the process of acquiring and appraising videos for long-term retention. It also includes all the activities related to providing access to archived video.

‘Preservation’ refers to the series of managed activities that ensures ongoing access to videos, with a focus on guarding against deterioration and technological obsolescence over time.

‘Management’ is the central piece, relating to ingesting, storing, organising, describing, locating and controlling video assets. This is a core and overlapping process that underpins and supports all other processes.

The project fully anticipated that no single application would be able to meet
the diverse range of campus needs. When exploring solutions, including the evaluation of a DAM system, based on the requirements of the entire video assets life cycle, however, a number of interesting issues arose. As this paper discusses, these issues essentially point to the practices of different communities, who despite all playing a role in the life cycle of video assets, traditionally do not work together. This parallel or siloed approach is also reflected in the software offered to and used by the communities, which tend to be self-contained, end-to-end systems, duplicating the core management function, and including tightly bundled specialisations to support one or more of the surrounding (or peripheral) processes described earlier.

**TERMINOLOGY**

Terminology is a challenge when working with diverse stakeholders and communities. The same terms can be used to mean different things and vice versa, causing confusion and miscommunication. The terminology barrier applies to different levels, ranging from essential concepts or definitions to detailed ways of naming things.

**Digital assets and DAM software**

In the context of UND, the term ‘digital assets management’ is used to refer to the activities related to managing digital information. The use of the word ‘assets’ implies an underlying process of making value judgments. Regardless of the formats, digital assets are records, data and resources, typically owned by the university, regarded as having (usable or historical) value and therefore needing to be maintained for a period of time and/or preserved in perpetuity. Three broad, not necessarily mutually exclusive categories have been identified to scope and conceptualise the digital assets landscape at UND:

- **University records**: The university’s ‘Records Management and Archives Policy’ defines university records as ‘recorded information created or received in the course of conducting university business...
and kept as evidence of such activity. This definition extends to records in all formats, including but not limited to e-mail, electronic and paper documents, film and print graphics, and audio and video recordings.

- **Research data:** The university’s ‘Data Retention and Access Policy’
  defines research data as ‘information recorded or customarily recorded in the relevant field, as a result of research. Data include notes, records, slides, photographs, drawings, information stored in electronic and/or computer readable form, reports, publications, correspondence, and summaries, compilations or derivatives of other data’.

- **Resources for teaching, learning and research:** This category of digital assets refers to the wide range of content the university collects, creates, receives as donation, or purchases to support learning, teaching and research. Examples include the library collections, donated datasets, electronic databases and electronic journals, as well as lecture captures and massive open online courses.

When evaluating DAM software, however, it is important to recognise that solutions only support the management of a limited subset that fits into UND’s broad definition of digital assets. As a matter of fact, there is a much narrower focus than the collective name suggests. Searching the web for a definition and types of DAM systems paints a confusing picture. Some define (digital) asset as ‘content + metadata’; content being media files that ‘have been classified, indexed, versioned, secured, stored, possibly reformatted or canonicalised in some way, and (typically) assigned a life-cycle state, a unique ID and an owner. These are the things that make a piece of content an asset’. Others see DAM systems as ‘a central repository’ for ‘the management, organisation and distribution of digital assets’ or more broadly ‘the collection of systems and processes that enable the management of assets from creation through distribution and archive’.

DAM systems are sometimes classified by their hosting options (on-premise or via software as a service) or the openness of the code base (proprietary or open source). Other classifications focus on the business functions or use cases being supported, such as brand management, marketing, publishing or video production. The broadest or most encompassing even include enterprise content management and digital content management as sub-categories.

DAM vendors seem to have picked a very broad term for a relatively small software products segment. There are different variants of DAM systems but they are difficult to differentiate clearly, making it far from straightforward for organisations to select the right product.

DAM systems are a breed of software that manages specific types of digital information within a specific organisational context. DAM is mostly intended for multimedia or rich media, such as photographs, videos, animation, graphics, logos and marketing collateral. DAM systems emerged in the private sector to support digital media creation, marketing, publishing and brand management, and their customer-base mainly consists of commercial organisations.

To date, the DAM Foundation’s ‘DAM maturity model’ and ‘10 core characteristics of a digital asset management system’ represent the best attempts to benchmark the industry and establish standards and best practices. The former is a self-assessment tool for measuring an organisation’s effectiveness in DAM, while the latter lists the essential functions or core competences that a DAM system must meet — any of which fit into the ‘management’ function that is central to and embedded throughout the video assets life cycle described previously. These criteria can still be used to assess and compare different DAM systems without having to
decipher the confusing self-declared or assigned ‘top’ position in various charts or reviews. The DAM Foundation also had a certification programme, and published a list of accredited vendors based on the core competence. Unfortunately, the DAM Foundation became inactive and closed down in early 2017.15

Archiving and archiving
Commonality has been recognised between DAM and other information management professionals such as digital librarians, digital archivists and digital curators. Expertise from libraries, archives and museums (LAM) is thought to be relevant, transferrable and critical in structuring, describing and providing access to digital assets. Debora Fanslow covered this topic extensively in her brilliant four-part analysis entitled ‘Who needs a DAM librarian’.16,17,18,19 Fanslow places DAM within the larger context of information management, and maps out the skills of information management professionals against the competencies required for DAM, to show the implicit value they can bring to DAM as practised in the commercial sector, concluding ‘who doesn’t need a DAM librarian?’.

The convergence between the DAM and LAM communities is not only evident in the advocacy by DAM consultants and practitioners on behalf of librarians, but also in the presence of LAM professionals at DAM conferences, as well as dedicated courses on DAM as part of many library and information science programmes.20 Some businesses now employ ‘digital asset librarians’ and ‘digital archivists’ to oversee digital assets and take overall responsibility for DAM systems.

DAM vendors have already been taking advantage of LAM practices, albeit not necessarily in a conscious or explicit way. Terms originated in the LAM disciplines, such as catalogue, collection, gallery, metadata, index, library and archive, are frequently used and built into DAM software. In the context of a terminology discussion, however, a word of caution must be voiced, as these concepts may not necessarily carry the original or same meaning in the DAM world. For example, when the name ‘DAM Library of Photo Archive’ is used for a category of DAM systems,21 the mixed metaphoric references to libraries and archives make it plainly difficult to envisage what this might entail.

The different meaning or understanding of ‘archive’ (both as noun and verb, singular and plural), one of the essential terminology differences discovered by the project, may be the underlying reason for the creation of a ‘library of archive’.

The Society of American Archivists uses Laura Millar’s words to define ‘Archives’ and makes a distinction between archives (in lower case) and Archives (with capital A): the former referring to permanently valuable records while the latter is an organisation dedicated to preserving the documentary heritage of a particular group.22

In DAM systems, archiving is the process of identifying inactive files and moving them from one tier of the storage system to another, typically from smaller, more expensive spinning disk storage to cheaper tape storage of much larger capacity. Cost-effectiveness is the main driver for archiving, or more accurately ‘archiving to tape’. Despite being less responsive, it is possible to store more data on tapes, costing much less. Thus, infrequently used files end up in the ‘(tape) archive’.

A reasonable guess of the previously mentioned ‘DAM Library of Photo Archive’ is that this is a collection of inactive photos, possibly stored on tape, with a DAM component sitting on top and providing library functions such as metadata and searchability.

The University Archives is the dedicated campus entity responsible for collecting, maintaining and preserving the official
records of UND. Some of these records are now in the form of digital videos, photographs and audio files, which have enduring value and need to be preserved for the long term, together with their context and inter-relationships. Long-term Archiving (with a capital A) comprises a series of activities including appraisal, accessioning, processing, preservation and providing access (to archived records) and is not to be confused with ‘archiving to tape’.

**DIGITAL PRESERVATION**

Every information management discipline is grappling with the challenges brought about by digital information technology. As the corporate sector developed and practised DAM, the field of digital preservation also emerged and evolved in the last 20 years, with strong involvement from LAM and academia.

While DAM vendors make use of LAM-related practices as a framework to organise the content of DAM systems, these tend to be from the pre-digital or analogue practice. The development related to digital preservation seems to have escaped DAM vendors’ attention. Occasionally, a lone voice would question the lack of involvement by the DAM and content management vendor community in this work, but DAM systems and content management systems in general are designed without functionalities intended to ensure the long-term accessibility of digital assets. No doubt the corporate sector maintains digital assets for very different purposes (eg monetisation) and, unlike heritage institutions such as LAM, rarely needs to commit to keep anything in perpetuity. But what is the intended life cycle for those communication artefacts in DAM systems? Taking into account the rapid rate of technological change, it is hard to imagine that obsolescence is not a concern. What happens when a DAM vendor stops supporting older video codecs?

DAM vendors respond to their customers’ requirements. When creators and users of digital assets tend not to think beyond keeping assets (on tape) just in case they need them again some day, it is unreasonable to expect features in DAM systems that support archiving and long-term preservation.

There are a number of possible reasons why preservation of digital assets is not a mainstream concern:

- Twenty years is not long enough for common born-digital (ie natively file-based) media content, especially those without dependency on physical formats or carriers such as CDs and DVDs, to become obsolete or inaccessible at such a scale that would cause panic or be considered a catastrophic failure. Even if some things got lost, they are still trivial in the greater scale of digital assets.

- Digital assets for marketing and communication purposes are often created for the short term. Once the business value has been outlived, it seems reasonable to put away or even discard the inactive content after the associated campaign has ended.

Or perhaps this is again a terminology issue in that digital preservation is addressed in DAM systems but just not labelled so? Many DAM systems do contain inexplicit, overlapping functionalities, albeit meant to support different use cases, which typically can also be found in a digital preservation system. Below are some examples.

**Checksum**

A checksum is a calculated string of fixed length associated with a piece of stored or transmitted digital data, as the result of running a hash algorithm. Comparing checksums before and after data transmission or at regular intervals is a common method to detect error and ensure bit-level data...
Checksums can be generated automatically or manually in many DAM systems but the main use case seems to be de-duplicating files, rather than ensuring ongoing data integrity.25

Metadata
Many DAM systems automatically extract rich (technical) metadata from the media files and are flexible in customisation of metadata fields. This allows for the implementation of any standard metadata schema but unlike a Digital preservation system, DAM systems do not provide any built-in templates of common metadata schemas such as Encoded Archival Descriptor (EAD), Dublin Core, METS or MODS.

File formats
File formats define the internal structure and encoding of digital objects. They become obsolete when the ability is lost to interpret and render them in human-accessible manner, due to dependence on technology used to create the digital objects. File formats play an important role in digital preservation as they allow characterisation of digital assets and the assessment of digital preservation risks. Digital preservation systems make use of authoritative file format registries and include functionalities such as format identification and validation. They also support the workflows related to format migration, including the necessary verification.

DAM systems can interpret common media file types, provide access to, and use file format information for sorting, browsing and searching digital assets. DAM systems also support transcoding, which involves decoding and re-encoding already encoded audio or video files, changing frame size, bit-rate, codec or audio signal. The main use case for transcoding is to turn large audio/video files to codecs/formats most suitable for users’ or viewers’ bandwidth and devices, often using lossy compression and meaning original quality cannot be recovered — something considered unsuitable for archiving and preservation.

SOURCE OF TRUTH
‘Single source of truth’ has been cited as an essential benefit of DAM systems. Regardless of whether this means storing every data element just once or for a single DAM system to act as a central repository, achieving it in a higher education institution environment is extremely challenging. This has to do with the diversity of institutional activities, devolved or distribution organisational structure, but also how software in general is designed and built.

Mapping stakeholders’ roles to the full life cycle of video assets (Figure 1) makes clear that each group typically performs one or more processes without getting involved in all life-cycle activities, and the different groups typically do not work together. This pattern also applies to software — there are multiple, and purpose-built software applications in use for almost each stage of the life cycle. They tend to be self-contained, end-to-end systems, often based on very different technologies. Under the hood, however, each of them requires a core management function, relating to ingesting, storing, organising, describing, locating and controlling video assets, which is often duplicated in each of these monolithic systems. As a result, in order to get the tightly bundled specialisations that support one or more of the surrounding (or peripheral) processes, organisations pay multiple times for the core functionalities and end up with separate systems.

Any system that specialises in managing digital content, regardless if it is media, documents, web content or learning material, should take digital assets life-cycle requirements into account and support the peripheral processes. Many DAM systems are moving in the right direction. Some
integrate non-linear editing tools, some support lecture capture and publication to learning management systems or social networks, some offer hosting options and allow end users to access digital media. The main gap, however, seems to be archiving and preservation — the processes vital to the longevity of digital assets.

Moving things to tape storage, even with regular monitoring and auditing does not deal with the risk of technological obsolescence. How can DAM systems be the truth of source without ensuring the ongoing availability of digital assets? If they are only destined to deal with the current or usable value of digital assets, how should they interact with archives and digital preservation systems, which are considered sources of truth over time? What are the things that need to happen while digital assets are still in active use to meet the downstream archiving and preservation requirements?

Life-cycle support for digital assets is not an advocacy for DAM systems or digital preservation systems to include more functionalities or become bigger monolithic systems. This is a modular approach instead, whereby the system of solution can be decomposed into a number of components that can be mixed and matched in a variety of configurations. Storage and interface or presentation layers, for example, are components, and as it is commonly accepted, should be separated from other system logics. Perhaps the modularisation can go further and separate out the core assets management function, with self-sufficient modules that specialise in one or more peripheral processes, capable of connecting, interacting and exchanging resources with the core logic, and, if required, with each other.

Authoritativeness or truth is embedded in the many types of digital assets possessed or stewarded by the university, both current and historical. For this reason, there is no single source of truth, just as no single system exists that will meet all requirements or manage all digital assets (as broadly defined by UND). One can, however, achieve aggregated truth, from many sources. In this context, ‘digital object identification’, ‘open architecture’, ‘open standards’ and ‘application programming interfaces’ become highly relevant. It would be desirable to see more support in DAM systems for archiving and preservation — an obvious thing to do would be to extend the existing functionalities related to checksum, metadata and file formats, which drive the main digital preservation processes, or ideally archiving and preservation modules, which can be plugged in and work on top of different digital assets and documents management systems, regardless of the formats. This would offer the best of the DAM and LAM worlds and provide true life-cycle support for digital assets.

References
13. DAM Maturity Model, ref. 10 above.