Turning information professionals into digital archivists for success in today's world of digital asset management

Received (in revised form): 28th March, 2013

Rob Corrao

is Chief Operating Officer of LAC Group, responsible for hands-on leadership and oversight of business operations and development across LAC's service lines. He has over 20 years of experience in information technology, strategy, operations, consulting, sales and marketing. Rob is well versed in the information technology needs of both traditional and special libraries and an active member of his community — he serves on several advisory councils, is involved with non-profit organisations, and has been a guest speaker at a number of digital asset management conferences. Rob graduated from the University of Notre Dame with degrees in both marketing and psychology, with concentrations in finance and Italian.

Tel : +1 323 852 1083; Fax +1 323 852 1093; E-mail: rcorrao@lac-group.com

Phil Spiegel

is Director, DAM Projects, LAC Group. Phil is a media archive professional with over 20 years of experience. He has managed large commercial and broadcast media archives and is fluent in the technical, operational and financial challenges in their efforts to digitise, catalogue, preserve and monetise content. Phil has also worked at National Geographic, Corbis Motion and Getty Images. He is active in the professional community, and along with blogging about digital asset management and media archives, has panelled and presented at numerous related events.

LAC Group, 6500 Wilshire Blvd, Suite 2240, Los Angeles, CA 90048, USA
Tel : +1 323 852 1083; Fax +1 323 852 1093

Abstract The subject of training is often undervalued, but it is the key to 'physical to digital transformation'. The traditional library/archive is fast becoming a thing of the past, and the challenge is to redefine the librarian's relevance and apply their information management experience to the new digital world. Today's digital archivists must possess the special skill-set necessary to organise, preserve and protect assets, but they actually add value to those assets by creating a link between the assets and their use. Training for digital archivists must include managing documents and images (moving and still), audio digital archiving with intelligent metadata tagging, cataloguing, indexing, digital file formats and delivery. This paper describes how digital archivists can work smarter with smart workflow, the importance of metadata for digital assets, intelligent cataloguing and business taxonomy, and includes a case study of a project undertaken by LAC Group.

KEYWORDS: cataloguing, archivist, library, metadata, managed services, professional services, research training, transition, taxonomy

INTRODUCTION

The subject of training is often undervalued, but it is the key to 'physical to digital transformation', and critical to the management and monetisation of assets — physical, digitised and

digital-born. In today's world of digital asset management (DAM), a combination of organisational, technical and intuitive skills is required to transform information professionals into effective digital archivists.

The traditional library/archive has become, or is fast becoming, a thing of the past, and the challenge is to redefine the librarian's relevance, expand their technical skills and apply their information management experience and good practice to this new digital world.

WHAT IS A DIGITAL ARCHIVIST?

Although digital archivist positions are relatively new, they have quickly become a necessity in almost every industry. According to *The New York Times*, there are over 20,000 people working as digital archivists throughout the world in every industry today, and that number is projected to triple within the next decade.[1] Whether one works in corporate America, the private business sector, government or law, a good portion of business is now conducted in the virtual marketplace, and that volume is expected only to increase over time. Therefore, the librarian comes out from behind the reference desk and becomes the digital archivist: an active participant, along with other institutional leaders, in defining the deposition of assets, their storage, access and reuse. Given this wider scope and new environment, the training, hiring or outsourcing of a competent digital archivist, undertaking some or all of the following, is crucial for future success:

- the organisation of digitised assets (associations, cataloguing, indexing, meta-tagging, etc);
- ongoing digitisation and integration of new assets, both physical and digital-born, into an existing digital library;
- digital and physical archive, file and storage management;
- digital asset use and retrieval;
- recovery of archives, assets lost in a crisis (or on the verge of loss), storage and expansion;
- client training on retrieval and use of digital assets to increase or monetise their value;
- permanent ongoing administration of the digital asset library;
- providing leadership on policy and governance.

Today's digital archivists must possess the special skill-set necessary to organise, preserve, manage and protect assets, but, more importantly, they actually add value to those assets because digital archivists create the link between assets and their use. Information professionals who possess only a strong IT background often lack the depth of knowledge required to understand how assets are used from a business application perspective. Traditional librarians sometimes lack the technical expertise needed to copyright, catalogue, tag metadata and organise assets in digital format for online access. Today's information professionals must have both technical expertise and strong organisational proficiency, along with intuitive business knowledge to understand how the customer (internal, external or both) uses those assets.

The unique cross-pollination of these two tracks and the resulting hybrid entity, the digital archivist, is the next step in the evolutionary process from the physical to digital. This convergence transforms and expands the role of the librarian from being that of a gatekeeper and a steward of physical assets to that of digital enabler, who manages, processes and curates digital assets, bringing opportunity and transparency into the new virtual library. Further distance from the old model is made daily as more and more assets are

born digital and live exclusively in this space, without ever having a physical representation.

Digital archivists need to be trained on how to handle copyright and licence challenges properly; to advocate and provide guidance on obtaining proper clearances; and to assure proper reuse and limited liability. Additionally, they must administrate various database tools, and evaluate, purchase, integrate and manage digital and online content from vendors, including online marketing tools. They assign identification markers, known as metadata tags, within the electronic data, to catalogue electronic material properly for high-demand access, and even manage numerous iterations of assets as data are migrated or new releases, versions and updates are applied to the DAM eco-systems managing the digital assets.

Training for today's digital archivists must also include managing documents, images (moving and still), and audio digital archiving with intelligent metadata tagging, cataloguing, indexing, digital file formats and delivery. Strong organisational and research skills are needed to handle complex queries and support user access. Digital archivists need to understand the digitisation process, content management, metadata management, content priorities, workflows, and also be knowledgeable about data storage and IT to help manage the secure digital environment where the assets reside. These professionals are expanding their abilities to include knowledge management, copyright management, competitive intelligence and much more. It is essential that digital archivists have a strong working knowledge of information technology and office technology applications, as well as familiarity with asset management, captioning and metadata management.

Familiarity with electronic databases and tools such as Factiva, LexisNexis, Bloomberg, Factset, Thomson ONE Analytics, etc is often also a requirement. Modern digital archivists need to be trained with the technical skills necessary to function as administrators and curators of the new digital library, providing different users with different access levels to assets, as needed. Finally, digital archivists must possess tactical skills at the enterprise level so they can train and support others in the retrieval of/access to digital assets in order to realise the value of those digital assets.

There is an increase in the level of complexity and usually a corresponding (though not always equal) reduction in overall staffing, as these new roles are able to work smarter and more efficiently. However, it is clear that digital archivists are a key component in the successful implementation and ongoing operation of a DAM system.

DIGITAL ARCHIVISTS WORKING SMARTER

Strategy and workflow

There is a long, rich history of library and information management and good practice for new digital archivists to draw from, adapt and apply to digital archive management, strategy and workflow. Organisations understand the opportunities and operational benefits that can come from digitised assets and a digital workflow. The efficiencies these processes can provide create higher processing and throughput rates without increasing staff numbers, as well as generating more time for these highly-skilled resources to apply sound business, content and archive strategy to the oceans of content that organisations manage. Digital archivists, as a part of their natural evolution to content curators, should have the bandwidth to address strategic content and collection issues, ie the 'what' about the archives, not just the

'where'. This is one of the great benefits of the transformation to a digital environment and why DAM is spoken of as an enabling tool.

Smart workflow eliminates process redundancies and, through user rights management, allows for disparate groups to process and develop information about digital assets simultaneously — adding to the cumulative metadata record set a variety of information about the asset, eg its provenance, technical information, physical asset data, usage rights and limitations, descriptive information, reference tags, location, etc. For example, once the core inventory record of an asset is defined, that information is never re-entered as the asset passes through further departments for rights review, metadata expansion, digitisation, etc. Each area is able to add department-specific information that enhances the overall record set, without unnecessary data entry or creating opportunities for error.

Digital archivists and DAM-based digital archives have to have a long view. The goal is secure, accessible and long-lasting viability of the digital assets and the support of both current and yet to be discovered uses for these assets over time. Digital archivists shape the sound archive management policies needed to support these ongoing efforts.

Process

Disparate groups, possibly in different locations, may be working on different parts of the physical-to-digital workflow and production pipeline. Any lack of unity and cohesive oversight means that this separation can lead to variances in the workflow and metadata policies for those working within these silos, and may ultimately undermine the full value of the DAM effort and the potency of the digitised assets. This can be a problem in any environment and especially so in a library, archive and DAM environment.

The digital archivist becomes the overarching intelligence, ensuring unity and assuring consistency for all assets admitted to the digital collection. Typically, the concerns of the digital archivist, as a part of this oversight, include:

- Does the output from all processes and the results still meet all the requirements of process and workflow?
- Are metadata requirements being consistently met?
- How do changes in the production and user cycle affect how/why/if at all certain tasks are performed or prioritised?
- Is workflow maximised for peak efficiency and effectiveness and, most importantly, is this happening while respecting the intellectual and physical integrity of the original materials? If not, what has been learned from the processes that are under way, to allow further refinement or (as needed) course correction? Is there a better, smarter way to do what needs to be done, and are there adjustments to the workflow to address?
- Is there good communication across the organisation, creating opportunities to share knowledge and allowing everyone to be effective and to feel empowered to participate fully in the transition from physical to digital and the success of the DAM project?

A digital archive workflow makes it possible to be better at managing metadata and assets to avoid unnecessary processing. The data record is enhanced without redundant data entry, with relevant pieces of information being added at each stage of the digital workflow — from shoot, to capture, to content, to rights, etc. The data snowball grows via a robust process into a complete archive, with its potential value increasing at each stage of the process.

As previously mentioned, DAM is an enabling tool. It removes barriers to access and creates opportunity. In an archive environment, it has the added benefits of fuelling access to, and further interest in preservation efforts to illuminate fully 'dark' archives (archives that are not digitised and accessible to the broader user community) in ways that add value far beyond the monetary. For example, digital archivists are ideally trained and positioned to lead content strategy and prioritisation efforts that will be in sync with known usage patterns. They understand what people want from their archives.

Content

Not all content is created equally. Some assets are more important than others. That importance is defined by the digital archives mission and/or the business rules around the process, and can be driven by rights, shelf-life or timeliness, format, location, and the topic and context of the content, etc. Digital archivists need to be content managers, driven to liberate content and allowing for access, reference, re-use and an otherwise long and meaningful life cycle.

Digital archivists, as stewards of the digital archive collection, will need to be trained to be content strategists who can effectively triage a library collection, prioritise and ultimately be comfortable with the idea that perhaps not everything can or should be processed. It is not about the quantity *per se*, but the quality and relevance of the content. Content strategy and collection management has to take a bigger role in planning for the digital archive and the digital archivist needs to be leading this conversation.

DAM allows users to access the content. They get to shape their own experience and, via desktop digital access, get to define the transparency of the archive. The role of the digital archivist is to engineer and then support an environment that lets the users experience and, in turn, define the value.

THE IMPORTANCE OF METADATA ON DIGITAL ASSETS

Dating back to 1876, with the creation of the first standard proprietary library classification system, 'metadata' has been a business term closely associated with cataloguing information and assets. The Dewey Decimal System was built on a 3-inch × 5-inch (7.6cm × 12.7cm) card catalogue system containing the book title, author, subject and a short synopsis, and it created a numeric identification to direct readers to the section and shelf in the library where the physical book was located. The information displayed on the index cards in a school or community library is actually the metadata for the library's assets. Today, 'metadata' is not just a common term used for library assets, but a term heard in almost every industry as companies expand their presence and assets into the virtual marketplace. Specifically, 'metadata' refers to 'data about the data', as well as 'data about the containers of data', also known as the structural metadata. 'Metadata' is an information set that includes the following details about individual assets:

- means of creation of the data;
- purpose of the data;
- author of the data;
- date and time of data creation;
- placement of the data on a computer network; and
- the standards used or followed.

Digital assets also include technical metadata and functional metadata, both equally important, while some technical data might include the camera and lens used, the stock, the specific software involved in the creation of the asset, etc. Descriptive metadata, on the other hand,

are about individuality of the data, the specifics about the data, the who, what, where, how, why, and so forth.

Much like a book placed on the wrong shelf in a library, the value of which is lost when the book cannot be located, digital assets can be lost completely if the metadata are not properly classified and recorded, and connected to the asset and its relationship to other related/derivative assets as well. Like the librarians of the past, today's digital archivists must be trained on the critical nature of metadata and its direct relationship to the value of each asset digitised. Training digital archivists on how to establish and properly build a digital library using various technologies is the key to building a strong foundation for digital assets and the libraries of the future.

Intelligent cataloguing is critical to asset value

MARC (Machine-Readable Cataloging) standards are digital formats developed by the US Library of Congress during the 1960s, in order to facilitate, create and disseminate standard cataloguing practices between libraries in North America. Since then, MARC formats, developed and enhanced through international cooperation, have been the standard for cataloguing a plethora of records both nationally and internationally. Today, MARC 21 standards are used to define records within assets, further breaking metadata out into referenceable partitions of usable information, and include the following three aspects of each asset record:

- Binary Record Structure (ISO 2709);
- Numeric Coded Field Designations within each record; and
- actual content of the record itself in terms of metadata transmission standard.

MARC 21 standard cataloguing of metadata adds more depth to the value of assets at the field level. By using common standards for cataloguing physical and digital assets with common syntax tags and field names, assets can be imported into any integrated library system or DAM system.

Cataloguing allows the asset to be retrieved at the field level by the use of intelligent keywords. Dynamic retrieval at the detailed field level can only be achieved through intelligent cataloguing by the digital archivist. Otherwise, assets are stored without any means or commands for meaningful retrieval. As is the case with a photography studio that may digitise thousands of photographs for customers to access online, specific descriptive keywords must be applied at the record level so that any number of different keyword searches will retrieve and promote each of their stock images. Today's digital archivists must be trained on the technology and standards, as well as the business content of each asset they catalogue, in order to expand search and retrieve capabilities.

BUSINESS TAXONOMY: THE KEY TO FINDABILITY AND USE

Business taxonomy is the hierarchical classification used to organise all physical, digital and conceptual business properties at an enterprise level for a business. It is a modern and dynamic approach for mapping and retrieving all of an organisation's information, and includes tangible information like documents, digital assets and products, as well as unstructured data including business processes, workflow, human resources and customer relations. Business taxonomy adds intelligence and precision to data mining with contextual filtering applications, pushing/pulling content and delivering information with a results-oriented methodology. Most importantly, taxonomy addresses content

accessibility by developing business concept structure and organisation around the digital archiving and retrieval process.

By building a business classification system around an entire metadata schema along with intelligent cataloguing, taxonomy serves as the authoritative source to connect assets with the content management and content-driven applications being used, and more importantly it positions the digital archivist to respond automatically and accurately to the specific requests and needs of customers in real-time. In simple terms, business taxonomy further expands the findability and usability of assets with another layer of business intelligence. Digital archivists must be trained to understand the rules of engagement for each of the assets they archive, metadata tag and keyword classify for a specific business within a specific industry. This enables faster, more accurate searches and retrieval of data by staff, customers and prospects alike, which in turn gives data value in the modern world of digital enterprise. Although the most challenging aspect of training and ongoing training, this level of intuitive knowledge is ultimately what transforms the librarians and information professionals from the past into successful digital archivists of the future.

CASE STUDY

Managing all aspects of the physical archive transition to complete digital archive at a major entertainment content creator

LAC Group is currently engaged in facilitating the transition of a decades-old physical media archive (consisting of related images, art, documents, scores, posters, promotional materials and other ephemera) at a major entertainment client to a state-of-the-art digital archive repository, with the goal of completing an overall digital inventory, providing digital access and opportunity for their client base, and identifying critical 'at risk' assets.

LAC Group was entrusted with this project, which includes creating a complete asset-by-asset inventory, building new data records for each asset, doing a condition assessment and then a content triage to prioritise digitisation in sync with the ongoing business needs of the parent company. LAC Group was able to demonstrate how to achieve an ongoing strategic advantage over competitors for digital access and operating cost containment, and how to drive opportunities for content mining and monetisation beyond the limited scope the physical archive provides.

LAC Group has a demonstrated speciality in physical and digital asset, library and content management. The group has been leading the executive-level conversation on strategy, operations, workflow, and approach for increasing digital access, physical to digital migration and on increasing content visibility and monetisation opportunities. The client has been able to tap into the LAC Group expertise to appreciate lower costs as well as improved service, training and support for their digital archive.

As a part of this overall project, LAC Group is helping the client evangelise on the benefits of this transition and is assisting with the change management, staff training and knowledge transfers needed to assure a complete transition. This transition is not solely for the assets, and also involves a complete transformation and training of their library archive staff and the supporting workforce as digital archivists, cataloguers and digitisation technicians.

Project details

LAC Group archivists are trawling through millions of assets, doing a complete

inventory, reviewing the data records and performing an asset review and content cull, so as to prioritise a subset of assets, within the overall collection, for digitisation and inclusion in the client's DAM environment.

The team has split tasks to create the most effective workflow to achieve the client goals and timelines. Experienced media archivists are reviewing all assets, performing the selection cull and flagging assets of significant historical or corporate importance — especially those items that collectively represent icon images and assets, where this particular media entity is considered.

As part of the review, assets are properly re-sleeved and re-packaged for long-term and stable archival storage. Items queued for digitisation are segregated, prepped for scan, and their data records are updated. Any information on or within the asset is captured, including locations, talent, year and other significant metadata. Records are made as complete as possible with the known information.

Further efforts are made to record the current state of the respective asset and to flag any items to be repaired, or needing any further digital cleanup or repair.

The scanning process creates archival quality TIFF files along with a corresponding proxy reference and data record for delivery to the DAM storage environment. Assets are then returned to archive and boxed for long-term, offsite, low-cost, archival quality storage.

LAC Group is also working with the user community and the client's internal library resources, providing training on the DAM tools and the overall workflow. This is to ensure that library staff members become digital archivists with the necessary skills to yield maximum value from the digital archive collection, and the ability to provide ongoing support beyond the scope of the initial project. There are many challenges, including:

- some physical or digital assets not properly stored;
- limited data to work from (in some cases), which leads to the need to manage expectations and workflow, and the need to extract or pull information out of the image/video — subject matter expertise;
- duplicates/non-original work mixed;
- non-owned assets mixed;
- not all assets accounted for in inventory (many inventories are done as time or resources allow);
- mixed media formats;
- assets spread across multiple storage locations;

There are also opportunities and benefits, including:

- complete and standardised inventory record for each asset;
- assessment of asset condition;
- content prioritisation;
- de-duplication;
- owned/non-owned assessment and sort;
- preservation of assets in jeopardy;
- opportunity to monetise best of collection; and
- proper archival storage/housing for assets not being digitised.

CONCLUSION

Digital archives are a reality and the bottom line is that DAM is an easy idea tied around a complicated solution. It is critical to the success of a DAM project to get the right people involved. Getting librarians and relevant non-librarians within an organisation properly trained to transform them into digital archivists to meet the challenges of this process will be a major factor in assuring the overall success of these archive transitions.

> 'Technology is nothing. What's important is that you have a faith in people, that they're

basically good and smart, and if you give them tools, they'll do wonderful things with them.'

Steve Jobs

Reference

1. De Aenille, C. (2009) 'Digital Archivists, Now in Demand', *New York Times*, 7th February, available at: http://www.nytimes.com/2009/02/08/jobs/08starts.html?_r=1 (accessed 24th May, 2013).

Useful links

DAM Ideas (various dates), DAM industry blog focused on the challenges of media archives and their transition to a digital asset management, available at: solution.http://spiegelams.typepad.com/spiegel_ams/

Earley & Associates (2012) 'What is the business value of taxonomy?', available at: http://info.earley.com/what-is-the-business-value-of-taxonomy/

'Presentation from the Digital Archivist at the Kennedy Library' 25th April, 2011):, available at: http://www2.archivists.org/sites/all/files/desc_sec_presentation_2011_boudreau.pdf

Smith, K. R. (2013) 'Defining the role of digital archivist', MIT Libraries, available at: http://libraries.mit.edu/sites/digital-archives/defining-role/

Society of American Archivists (Training Curriculum and Certification), available at: http://www2.archivists.org/