



What's Your Product?

Assessing the suitability of a
More Product, Less Process methodology
for processing audiovisual collections

by joshua ranger
senior consultant
audiovisual preservation solutions

What's Your Product?

Abstract

The widely referenced and adopted More Product, Less Process methodology (MPLP) represents a much needed evolution in the manner of processing archival collections in order to overcome backlogs and resource shortfalls that institutions face. **In the case of audiovisual-based collections, however, the ability to plan budgets, timelines, equipment needs, and other preservation plans that unequivocally impact access is directly tied to the documentation of some degree of item-level knowledge about one's collection.** This paper proposes an extension of the MPLP model which is necessitated to properly address the particular needs of audiovisual and other complex media in a way that properly meets archival standards and that assists the archivist in generating their true product: the provision of the three basic services of **Findability, Access, and Sustainability** regardless of the format, the content, or the tools used.

What's Your Product?

Introduction

In their ground shifting paper “More Product, Less Process: Revamping Traditional Archival Processing”¹, Mark A. Greene and Dennis Meissner outline the very real problem of backlogs in processing archival collections, recommending a new approach to processing that looks to minimize arrangement, description, and handling, all in order to minimize the time from ingest to providing some form of access to researchers. The exact impact of the widening adoption of a More Product, Less Process (MPLP) approach to archival processing may not be fully measurable for some years. What can be measured is the enthusiasm for the methodology, a fact that underscores the great need and desire for some kind of solution to existing (and ever increasing) processing backlogs.

What seems apparent as well is that the MPLP approach presents numerous holes in its application to audiovisual or other complex media collections that benefit more directly from item level documentation. There is a tacit confirmation on this in the fact that Greene and Meissner's original paper – titled in an early draft as dealing with “late 20th century collections” – makes no mention of audiovisual materials outside of the survey questions included as addenda. More explicit confirmation comes from the continued difficulties institutions are having with planning MPLP-style projects or meeting processing rates when dealing with audiovisual materials².

The issue here is not that there is no efficient method of processing large audiovisual collections. The issue is that MPLP-derived outcomes and metrics used to plan processing projects and measure their success *have not been and cannot be* extrapolated directly to audiovisual materials. First because the physical and intellectual makeup of audiovisual media does not correlate to those of textual materials. Secondly, unlike an MPLP-like approach that can successfully limit description to the collection or series level, **the ability to plan budgets, timelines, equipment needs, and other preservation plans that unequivocally impact access is directly tied to the documentation of some degree of item-level knowledge about one's collection.**

This kind of problem has often been the case when attempting to address media records with traditional archival/library science methodologies that were originally designed for print and text materials. A problem, to be sure, but not dire. However, if institutions with audiovisual collections (and this includes the majority of archives these days) are facing situations where budgets or grants are being doled out and staff are being assessed based on MPLP-related guidelines, we *are* at a point when such media type discrepancies will cause great harm to archival efforts.

¹ Accessed July 30, 2012 at <http://www.archivists.org/prof-education/pre-readings/IMPLP/AA68.2.MeissnerGreene.pdf>

² “Harvard University Libraries 2009-2010 Annual Report, Report of Marilyn Dunn, Executive Director of the Schlesinger Library and Librarian of the Radcliffe Institute.” Findings state that for audiovisual processing “Limited success indicates that additional staffing is necessary to keep this area from developing truly insurmountable backlogs.” and may also lead to acquisition limits. On-demand viewing copy creation has also been ended. Accessed 7/29/2012 at http://hul.harvard.edu/publications/ar0910/08-schlesinger_library.html

University of California Libraries “Survey Results & Analysis for Next Generation Technical Services: Special Collections/Archives.” Respondent states “While many Spec. Coll. departments focus on processing of linear feet, we work more along line of processing collections often with mixed media (audio, photos, etc.) Focus on linear feet is not really representative of good way to measure work productivity and potential for efficiency.” Accessed 7/29/2012 at http://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=11&ved=0CGIQFJAAOAO&url=http%3A%2F%2Flibraries.universityofcalifornia.edu%2Fabout%2Ffuls%2Fngts%2Fdocs%2Fngts3_spcoll-arch_survey_report_no_ID.DOC&ei=qX0VUNz3JuPo6gHM5YG4CA&usq=AFQjCNFQw8MzYFI78TRLWT9EeYm2sfunYg

The Smithsonian Archives of American Art are beginning a 3-year project that will in part “develop benchmarks and guidelines for an archival approach to processing and describing archival collections with audiovisual content” prompted by the perceived “gaps in current archival standards and best practices”. Uncovering Hidden Audiovisual Media Documenting Postmodern Art website accessed 7/31/2012 at <http://www.aaa.si.edu/collections/projects/clir>

Greene himself in a Q&A session stated that unlabeled audiovisual materials are typically not worth acquiring and “clog the shelves”. Simmons College GSLIScast “Mark Greene - On More Product, Less Process”. Accessed 7/30/2012 at http://gslis.simmons.edu/podcasts/podcast_extras/2009/20090415-greene-transcript.php

What's Your Product?

Ultimately, the cost and effort of archiving – as well the wide existence of important collections outside of traditional research libraries or similar institutions – means that the idealistic notions of preservation for preservation's sake or for the sake of a limited number of potential researchers are not a sufficient enough advocacy argument. **Instead, efforts at processing, collection management, and advocacy must be centered around supporting the goals and health of the archive and the parent institution as a whole.** For paper-based collections in university-type archives, MPLP may be considered sufficient for achieving the goals of drawing in researchers and providing access to those collections. For media collections in traditional archives and for many collections in non-traditional archives, MPLP *as-is* is neither sufficient to support archival and preservation activities nor sufficient to support the continued budgetary allotment for archives within those institutions.

Metrics & Outcomes

At their base the primary metrics and outcomes of MPLP include the measurement of processing rates in linear feet per [time unit], and the creation of a finding aid or other mechanism that allows researchers to find, request, and browse materials.

Though the use of linear feet can make sense when dealing with paper and photographic materials, it is an empty measurement for audiovisual materials. Formats of varying width and size are typically mixed on the shelf or within boxes. Items may be stacked multiple layers deep both on the horizontal and vertical planes, or in positions that fill in gaps or maximize space. Even if formats are not mixed, a box may fit only 15 U-matic videotapes or several hundred audiocassettes, each of which probably has a different duration. And this variation in duration among items of the same physical size -- the time-based nature of media -- is a major differentiation in asset size when moving to the digital realm.

In terms of intellectual arrangement, the relationship of object to content with audiovisual materials makes little sense to discuss in a folder/box arrangement measured in linear feet. A physical carrier may hold one or multiple works, a single work may be divided across multiple carriers, or the components of a work will likely exist across multiple carriers, which may also exist across multiple formats. Even if one identifies a linear foot as containing twelve 16mm motion picture film reels, that is a highly inaccurate representation of the intellectual arrangement and content of those reels. Defining the high-level content type categories (Raw Footage, Home Movies, Projection Prints, Oral Histories, etc.) is equally important to determining at what level to arrange, describe, and measure progress.

This then begins to touch on the desired outcomes of a MPLP process. Broadly speaking, yes, the desired outcome is greater researcher access. What enables that access is the assessment of the material's condition for handling and long term storage, along with the creation of a finding aid or other access point. Under MPLP, minimal processing and higher level finding aids then rely on researcher driven discovery and access – presenting enough information to guide a researcher to the likely set of folders or boxes and allowing them to browse the lesser described items within.

Audiovisual materials do not always fit well into such a researcher-driven setup. First, if items are un-described and (as is often the case) have few or no annotations to identify them, their content cannot be guessed without playback. Second is that very issue of playback. Many archives do not have the equipment to playback even a portion of the audio or video formats they hold, or do not wish to allow playback due to risk of damage from poor condition, machine failure, or user error. Some archives will perform on-demand reformatting to create an access copy, though often at a cost to the researcher and after a period of delay.

Needs & Approaches

On their face these can appear to be nitpicky issues – the continuation of sniping or complaining about the lack of respect afforded to audiovisual archiving needs. But these issues matter because they represent blockades to archives fulfilling their duty, impediments to distributing their “product”. **And at the crux of this issue is defining what the product in More Product is.** The adoption of the term asset in archival practice may confuse us into thinking that the manuscript, the film, the daguerreotype is the product, because in the non-archival world an asset or product is an object or idea with tradable value. Thinking beyond that stance, one might consider that the finding aid created via processing is the product. Neither is true. The archivist's product, as I define it, is provision of three basic services:

- Findability:** Materials should have sufficient physical or intellectual documentation or arrangement to provide basic capabilities of searching, browsing, and other methods of discovery that prompt access.
- Access or Usability:** Where rights and security provide, materials should be readable, viewable, listenable, copyable, functional in applicable hardware/software, distributable, etc. within a reasonable timeframe, to a degree that represents the appropriate level of quality and integrity of the original, and in a manner that supports the defined mission and needs of the organization.
- Sustainability:** Materials should maintain these aspects of findability and accessibility with optimum integrity and reliability across time, locations, platforms, systems, and changes in technology and ownership.

These services are the responsibility of the archivist regardless of the format, the content, or the tools (such as a finding aid) used to fulfill those goals.

What this means for audiovisual collections in summary is that content must be findable independent of annotations on the physical item; content must be accessible or easily made accessible no matter the format; and the content must be sustained over time across multiple formats and systems. What all that means is that one must have sufficient intellectual control over one's assets in order to plan for and provide the means for findability, access, and sustainability. This includes the ability to:

- Plan for acquiring or maintaining playback equipment for identified formats
- Plan for budgets and approaches to in-house or outsourced reformatting
- Plan for periodic reformatting/migration in the future
- Prioritize assets based on a combination of technical, usability, and research factors
- Plan for budgets and approaches to long term, preservation-quality storage for physical and digital assets
- Provide access points to materials of sufficient quality and via sufficient platforms in order to support meaningful use across multiple user groups

An important point to underscore here is the absolute necessity for reformatting when dealing with legacy audiovisual materials. **Without the ability to playback an original format or to move that content to a new format that is not at risk of poor condition and/or obsolescence issues, an audiovisual collection is effectively meaningless.** As stated earlier, collection management and preservation planning require some degree of item-level intellectual control over a collection. This is not necessarily at the level of content description, but at a level that captures format, approximate durations, content types, and other data points that feed into such planning needs. These needs can be addressed efficiently, but the approach must be at a deeper level and must be assessed by different metrics than are currently under consideration by administrators and funders focusing on MPLP as stipulation to greenlighting projects.

What's Your Product?

What is suggested is an outcomes-based approach – like MPLP – but one where the outcomes have been realigned to match with the particular needs of accessing and preserving audiovisual materials. Based on my experience performing collection assessments and inventories for a wide range of traditional and non-traditional audiovisual archives, I have developed a basic approach to addressing unprocessed collections I outline below. I should clarify that, from my point of view, ideally all audiovisual collections should be documented at the item level. Pragmatically, this is not always possible as a near term solution to jump starting preservation projects. The realities of resource availability versus work necessity must be considered, and in such cases I look to approaches that are Item-ish versus those that are Collection-ish.

Both approaches capture data near the item level, but differ mainly in the amount of attention paid to each item. At the Item-ish level one documents a confined set of data points for each individual item, keeping in mind that not all points are available, not all groups of items deserve full depth of data capture, and the scope of one's efforts do not necessarily feed into a finding aid or library catalog record. At the Collection-ish level one defines a set of potential categorizations of intellectual arrangement that can adapt to the shape of the collections and the goals of processing. This may be at the format level across all collections, at the format level within each subcollection or series, broken down even further within a collection/series/location/ etc. (i.e., each format grouped by content type or date range or other parameter), and so on. At this level one documents a much more prescribed set of data points that can equally refer to individual items as to large groups of items.

Outline to Developing an Audiovisual Processing Approach

Define your product:	What are the goals for a processing project and for the collection? What are your desired uses and who are your potential user groups?			
Potential Uses:	Access	Advocacy	Distribution	Planning/Budgets
	Reformatting	Fundraising	Outreach	Collection Management
Potential User Groups	Students/Researchers: Internal and external parties accessing content for research or utilizing it for educational purposes			
	Communications: Promotions, public affairs, development offices, journalists, etc. seeking content for production			
	Institutional History: Parties tracking or using the history of the institution in support of its missions and goals			
	Public: External parties purchasing/ accessing content or interacting with it through social media			
	New Content: Any party using or licensing existing content to create new work, or creating wholly new content that may become a part of your collections			

What's Your Product?

Define your Depth: Determine the depth of processing required to produce your Product – How much and what type of information will support your work post-processing? At this point it is more inventory than finding aid. It may be an item-ish level inventory (data captured for each item, though not necessarily the full set for each) or collection-ish (data captured for each item at a defined level of intellectual arrangement such as collection, series, media type, etc.).

Define your Data Needs: Think Short Term. Think Long Term. – What are the immediate needs for utilizing this data set? Do they merge with your long term data needs or stand outside of them? What use will this data serve in the future? Unless a part of your prioritization decision, do low level technical or descriptive points specific to a particular instantiation matter? If capturing such granular data points, does the data need to move ahead with the asset record in the future or can it be dropped post-reformatting?

Define your Data Set

Item-ish Level
Focus on technical data points, minimize descriptive data points (title and primary contributors), and review condition.

Prioritize your Focus
Spend more energy on high value materials and formats known to be problematic or that have a known problem storage history. Check condition in enough items to verify existence or absence of problems, then spot check afterwards.

Know when to Extrapolate
Leverage any potential arrangement of materials, existing data or collection knowledge, or knowledge of audiovisual production and materiality to extend or extrapolate data. If a group of materials presents the same format/duration/element type or similar values, focus on capturing unique values like title and programmatically add repeat values later.

Collection-ish Level
Develop a minimal set of data points to collect for each item or group of items within a collection/series. Consider Quantity, Format, Duration, Date Range, Content Type, and Condition. Arrange by format, by format and content type, by collection and format, etc.

Determine how detailed your breakdown will be. How broad of a date range will items be categorized in? How exact or approximate will durations be? How exact will condition assessment be (specific instances or general assessment)? Verify known problems. Low value or non-problem formats can be quantified with little item-level review.

Determine at what point you can estimate or assume like values for a group of assets, or at what degree inexactness balances out. Is it worth your time to identify ten 60-minute cassettes within 300 90-minute cassettes? For distribution or classroom planning those lengths may matter, but in the grand scheme of reformatting the difference is not huge. If the ultimate goal is reformatting, what degree of condition assessment is necessary for prioritization? In other words, if you are transferring regardless of findings, does condition matter except where impacts selection and workflows?

What's Your Product?

Choose your Tools

If you don't have a system in place yet or plans for acquiring a system, focus on tools that are utility-like, for capturing data and doing something else with it, not heavier catalog systems. If a schema is not in place, think modular and granular to break down fields in a way that can more easily be mapped into other standards without excessive data wrangling. Existing tools such as the METRO/AVPS developed AVCC³ toolset may be a good option.

Focus on lightweight (can you develop something quickly and independently), portability (can you get the data easily in and out to use it elsewhere), control (does the tool allow you to control certain input values), and calculation (does the tool allow you to quickly arrange and calculate totals without extensive data wrangling). For long term data management and discovery, a relational database, OPAC or finding aid makes sense, but for planning a granular spreadsheet or flat Filemaker database may be sufficient to quantify and analyze.

And Then: Make it happen.

Example of Collection-ish Data Set

Location:		Unit B, Shelf 2			Collection:		Regional Folklore Collection		
Quantity	Media Type	Format	Size	Base	Approx. Duration	Date Range	Content Type	Condition Note	
10	Sound	1/4" Open Reel Audio	7 inch	Acetate	30	1960s	Oral History	Brittleness and curling; some cinching	
35	Sound	1/4" Open Reel Audio	7 inch	Polyester	60	1975-1982	Oral History	Minor display of SBS in later reels	
45	Sound	Betamax	--	--	60	1985; 1960-1982 (content)	PCM-F1 transfers	none	

³AVCC (Audiovisual Community Cataloging) is a free set of guidelines, planning documents, and cataloging tools designed to help archives document audiovisual collections more efficiently while utilizing a short term volunteer effort. Its focus is on creating data that can be used for planning preservation projects, advocating for collections, and mapping information into existing or future database systems. AVCC is being developed in association with the Metropolitan New York Library Council (METRO) with a planned initial release in fall 2012. Watch <http://www.metro.org> for future news.

What's Your Product?

Example of Item-ish Data Set

Primary

Collection	Regional Folklore Collection
Location	Unit B, Shelf 2
ID	ABC123
Media Type	Sound
Format	1/4" Open Reel Audio
Title/ Description	Ole Oleson, "The Bandage Man"
Size	7 inch
Base	Acetate
Duration	30
Date	1960s
Commercial / Unique	Unique
Content Type	Oral History
Condition	Some curling at head
Speed	7.5ips
Tape Thickness	1.5mil
Track	Full track
Sound	Mono
Noise Reduction	None
Generation	Original
Part	1 of 1
Subject	folklore, Oregon, scary stories
Contributors	Ole Oleson; Jimmy Jameson
Copyright	Onsite use only

Secondary - Helpful for planning if available but not always readily apparent.

Tertiary - More likely to be captured post-reformatting. If available may help finer prioritization.

Conclusion

This outline to a processing approach should be considered more as principles of a methodology. The implementer will still need to make informed decisions on scope, content, and tools. What informs these decisions is, in part, archival practices, but the guiding principle is the outcome. What is the purpose of processing the collection? What is your goal for collecting and using the data? What is your product?

AudioVisual Preservation Solutions is a full service preservation and information management consulting firm. AVPS provides effective individualized solutions founded on our broad knowledge base and extensive experience in the areas of collection assessment, metadata management, research & development, training, workflows, facilities design, and more. With a strong focus on professional standards & best practices and the innovative use & development of technological resources, we aim to help our clients achieve efficient, high-quality capabilities to meet the challenges faced in the preservation and access of audiovisual content, digital assets, and institutional data.