Yale University Libraries

Audiovisual Collections Inventory Project Final Report

AVPreserve
253 36th Street, Suite C302
Brooklyn, NY 11232

josh@avpreserve.com

Telephone
917.475.9630

October 2015

Table of Contents

Introduction

Project Overview

During the summer of 2015, in collaboration with the Yale University Library Preservation Department, AVPreserve managed an inventory of unique audiovisual materials held by special collections within the Divinity Library, Medical Library, Arts Library, and Music Library. The Music Library materials were inclusive of the Historical Sound Recording collection, the Oral History of American Music collection, and the Music collection. The inventory covered materials stored both onsite and at the LSF offsite storage facility. Commercial materials were invariably mixed in with larger collections or made up the bulk of certain named/donated collections and were thus included in the inventory. However collections that were known to be only commercial in nature (such as the large 78 album collection maintained by HSR) were not included in the inventory. Exceptions were made for certain commercial collections where the type or source of the material was considered to be rare or nearly unique due to age or scarcity, such as the Berliner collection or Cylinder collection.

Eight student workers were employed for the duration of the project. The inventory consisted of two activities: The capture of photo documentation of each item followed by data entry completed based on those photographs. This approach was taken in order to limit the amount of movement of materials – most of the photography could occur directly at the shelves or in close proximity to where materials were stored and immediately returned – thus not burdening YUL staff with the need to pull and/or transport boxes and tapes. The photographs also provide added documentation to researchers that could not be captured in the timeline of the inventory, such as extended lists of performers, song titles, graphical annotations, and other annotations related to provenance and description. LSF-stored materials were not able to be photographed at the shelves. Instead, daily shipments of totes were delivered to the Music Library for photography and returned the following morning at the time of the next delivery.

Records for 47,096 distinct items were captured. A breakdown of formats and other parameters is included below.

Successes

- Project was completed within the projected timeline.
- Significant increase in intellectual control, especially for collections that previously had had only box level or collection level description.
- The increase in intellectual control was even greater in terms of materials stored at LSF and certain sub-collections that had been boxed at some point in the past with no box list and no institutional memory of the contents.

Challenges

YUL staff were helpful in completing the project, however a few locations presented challenges as far as ease of access or the quality of working conditions.

- Within certain sub-collections some boxes had been over-filled with materials, resulting in difficulty in pulling overweight boxes from higher shelves. A certain portion of these boxes were skipped in the inventory in order to avoid damaging materials and employee injury.
- Certain format types (open reel audio, film, lacquer discs) presented handling challenges due to fragility or poor storage conditions.
- Exact format identification could be difficult at points where physical characteristics are not distinguishing (such as Betamax and PCM-F1) or where the material was too fragile to fully inspect (such as determining the backing of some ¼ inch audio tape).

Summary of Outcomes

- The six repositories hold over 47,000 audiovisual items that are unique, rare, or a part of named collections. The majority of these items are audio recordings.

- The total number of items would be much greater if pre-identified commercial-only collections (such as the 78 RPM disc collection held by HSR) had been included. The addition of materials held by Archives & Manuscripts and the Beinecke would bring the total well over 100,000 items.

- There is a wide range of content types, including:
    - Student and faculty produced content
    - Recordings of campus events
    - Instructional materials
    - Personal and institutional papers donated/acquired by the repository
    - Oral Histories
    - Commercial and original recordings of research value to the repository

- Overall condition of materials is fair-to-good considering the age and formats. Attempts at playback may reveal issues that could not be determined at the level of the inventory, such as Soft Binder Syndrome and shrinkage.

- This does not mean that preservation reformatting is not a pressing need. Materials remain largely inaccessible in their current format or present a preservation risk if the original is accessed, the size of the collections means that current ad hoc or on-demand reformatting will not be able to address maximal reformatting within 5-15 years, and obsolescence and decay will occur sooner rather than later regardless of the quality of storage conditions.

- Intellectual prioritization will be a major component in the selection process. For most of the repositories this will be primarily in regards to the importance of content. For HSR there will be a greater degree of consideration given to best copy selection and the determination regarding the reformatting of commercial works, given that portions of the collection have been reformatted to open reel audio or CD in the past.

- Primary collection management issues to be addressed will be:
    - Rehousing of the HSR Test Pressing boxes stored in B20B.
    - Rehousing of items with unstable or no housing.
    - Improved wind (flat and even “library wind” with ends taped down and films on cores) for open reel audio/video and film materials in long-term storage. This will be a post-digitization outcome and does not necessarily need to be performed before that point, unless items are not slated for nearer term digitization or the winds are in such a state that the tape or film is being actively damaged.
    - Creation of comprehensive documentation accessible in a central location that describes where all materials are stored within the libraries (or offsite) and the general arrangement of materials on the shelving units.

Format Breakdown

Overall Counts

Audio

Format	Count
1 inch sound tape reel	16
1/2 inch sound tape reel	109
1/4 inch sound tape reel	302
1/4 inch sound tape reel - acetate	2,611
1/4 inch sound tape reel - paper	3
1/4 inch sound tape reel - polyester	14,876
8-track cartridge	4
Audio-CD	142
Compact cassette	8,524
Audograph	54
Berliner	665
CD-R	4,924
CD-RW	10
DAT	351
DTRS (Hi8)	2,301
Microcassette	56

MiniDisc	57
PCM-F1	14
Sound cylinder	440
Sound disc - 45	34
Sound disc - 78	423
Sound disc - Lacquer	2,714
Sound disc - LP	287
Sound disc - Other transcription	657
Sound wire reel	45
SoundScriber	146
Test Pressing	3,755
Total	**43,520**

Conclusions:

- The vast majority of audio materials are ¼ inch sound tape reel, primarily polyester-backed, followed by compact cassettes and then CD-Rs.

- As the largest collection, much of these materials come from the HSR collection, a not insignificant portion of which are derivatives or not unique (dubs from lacquer discs and 78s, commercial materials mixed in with unique items).

- A primary point of prioritization will be an intellectual assessment and decision on the reformatting of commercial works – are they considered rare enough to transfer for historical research purposes (such as the Berliners and sound cylinders), do they round out a collection that also includes unique materials, have they been reformatted and preserved elsewhere that can be a source for research needs, or have they been transferred by Yale previously in a format that is sufficient for preservation moving ahead and that may be easier to access for migration to file-based storage.

- From a technical standpoint, the highest risk materials are:
    - Lacquer sound discs: Though the discs at YUL were in good condition, the lacquer coating is highly unstable and can lead to total content loss once it begins to delaminate.
    - Paper-backed and acetate-backed ¼ inch sound tape: These backings are prone to high rates of decay and/or damage during playback, especially given their age.
    - Wire sound recordings: Very easy to damage during handling and playback. Certain metal types used are prone to corrosion.
    - DAT: Highly prone to failure in playback and of decks. A limited number of decks are available for transfer, and those are breaking down rapidly.
    - CD-R & CD-RW: Increasing reports of disc failure, even in items only 5-10 years old. Not trustworthy for extended storage of content.
    - 1 inch sound tape, ½ inch sound tape, PCM-F1, and DTRS are all specialized production formats. Proper playback equipment and expertise will be limited.

Video

Format	Count
1 inch video reel	9
1/2 inch video reel	7
2 inch quadruplex video reel	16
Betacam	194
Betacam SP	218
Betamax	2
Blu-Ray	1
CD-ROM	3
D2	11
Digital 8	2
Digital Betacam	2
DVCAM	1
DVD-R	169
DVD-RW	200
Hi8	4
MII	2
MiniDV	36
S-VHS	44
Umatic	653
Umatic SP	84
VCR VC-60	1
VHS	1,175
VHS-C	2
XDCAM	1
Total	**2,837**

Conclusions:

- More than half of the video materials are VHS and U-matic. Except among the VHS there is limited concern regarding the existence of commercially distributed materials among the collections. “Distributed” is distinct from materials that may have been produced for commercial purposes but represent a master recording or the outcome of the production process, which are more likely to be on formats other than VHS.

- Unlike the audio materials, there appears to be limited duplication or derivatives among the video. For any that do exist, it is likely that the transfer specifications do not meet current preservation levels and would need to be reperformed.

- The relatively low number of videos here suggests two prioritization approaches:
    - Perform very selective intellectual prioritization.

    - Prioritize based on technical risks (with some intellectual considerations) and digitize as much as possible.

- From a technical standpoint, the highest risk materials are:
    - 2 inch, 1 inch, ½ inch video: Among the oldest video formats and therefore at increased risk for decay. Playback decks are also becoming sparse.
    - D2 & MII: Short-lived professional formats. Few playback decks remain.
    - DVD-R & DVD-RW: As with CD-Rs, high risks for failure even for newer recordings.
    - Hi8, Digital 8, and MiniDV: Small sized formats that are problematic to transfer. Decks for the 8mm formats are becoming sparse.

Film

Format	Count
16mm film reel	253
16mm film reel acetate	48
16mm film reel polyester	137
16mm full coat mag track	6
35mm film reel	79
35mm film reel acetate	1
35mm film reel polyester	58
35mm full coat mag track	1
8mm film reel	16
Total	**599**

Conclusions:

- The majority of film materials are 16mm, which is typical of library collections that do not specialize in feature films.

- Due to the avoidance of handling film on cores outside of their container in order to prevent damage, the backing of many reels was unable to be determined.

- Outside of intellectual prioritization, a major consideration will be age and source of the materials. Older films will be more likely to be acetate (higher risk), and materials that may have been stored in poor conditions (especially in extremely hot and humid environments) will be at greater risk. Additionally, color film from the 1970s will be a risk. Commercial or educational/industrial projection prints will be a lower priority unless the content meets collecting policy.

Other

Format	Count
External Hard Drive	2

Film Strip	27
Floppy disc	5
Iomega REV Disk	55
Slide	23
Thumbdrive	1
Zip Disk	11
Total	**124**

Conclusions:

- Many of these disks and drives contain safety copies of digitally recorded audio, or the production files (such as from ProTools) for audio mastered from live recordings. As all formats are obsolete or at risk for data loss, an intellectual review is required to determine the relative value of the files and/or the existence of other copies.

- Some materials, such as the Thumbdrive, are a part of artworks or have additional three-dimensional, non-audiovisual objects related to them. Upon intellectual review, these may require an alternate approach to preservation along art conservation lines, but the fact remains that their content does need to be migrated and recreated/emulated in order to persist.

By Repository

Arts

Format	Count
1 inch sound tape reel	3
1/4 inch sound tape reel	2
1/4 inch sound tape reel - polyester	22
Audio-CD	15
CD-R	8
CD-ROM	2
Compact cassette	46
DVD-R	13
DVD-RW	1
Microcassette	4
Thumbdrive	1
Sound disc - LP	23
Umatic	9
VHS	75
Zip Disk	5
Total	**229**

Conclusions:

- The bulk of materials are VHS, compact cassette, and ¼ inch audio.
- There are three primary types of materials: MFA thesis materials, works or interviews/talks from individual donors, and commercial materials contained within individual donor collections.
- As mentioned above, there may be different preservation/conservation decisions to be made based on the content and formats: whether content can be transferred per normal specifications, whether content needs to be emulated or conserved/restored, and whether the commercial content owned by donors is significant enough for research that it should be digitized or if the mere existence/record of the physical object is sufficient.
- From a technical standpoint, the highest risk formats are:
  - Microcassette
  - Umatic
  - Any optical or digital storage media
  - Open reel audio and compact cassette, dependent on age and condition

Divinity

Format	Count
1/2 inch video reel	3
1/4 inch sound tape reel - acetate	126
1/4 inch sound tape reel - paper	3
1/4 inch sound tape reel - polyester	93
16mm film reel	103
16mm film reel acetate	3
16mm film reel polyester	6
35mm film reel	17
8mm film reel	15
Audio-CD	2
Betacam	184
Betacam SP	68
CD-R	320
CD-RW	8
Compact cassette	1,885
D2	11
DTRS (Hi8)	3
DVD-R	66
DVD-RW	1
Film Strip	27
Floppy disc	5
Hi8	4
Microcassette	34
MII	2
S-VHS	1

Slide	13
Sound disc - 45	2
Sound disc - 78	3
Sound disc - LP	20
Sound wire reel	1
Umatic	141
Umatic SP	82
VHS	521
VHS-C	2
Zip Disk	**4**
Total	**3,781**

Conclusions:

- The vast majority of materials are compact cassette and VHS, followed up by Umatic, CD-R, and open reel audio.
- This is one of the most varied collections according to format, which makes sense given that much of the material is sourced from international organizations and institutions (as well as internally).
- This diversity has its pros and cons. At a much lower cost impact, whole sets of at risk formats can easily be digitized. However, these may actually be a higher cost per item if not included in bulk reformatting discounts, and the very large groups of formats may require a greater degree of intellection prioritization.
- Complex works such as film strips and slide shows may require additional considerations – as with art works – regarding their post-digitization presentation and intellectual control.
- The highest risk formats include:
    - Sound wire
    - MII
    - ½ inch open reel video
    - Acetate-backed ¼ inch open reel audio
    - Optical media and data storage media
    - D2
    - Acetate-backed and early 20th century film.
    - Any film, open reel audio, compact cassette, Umatic, and VHS, dependent on age and storage history.

Medical

Format	Count
1 inch video reel	4
1/2 inch video reel	4

1/4 inch sound tape reel - polyester	2
16mm film reel	32
16mm film reel acetate	42
16mm film reel polyester	43
35mm film reel acetate	1
Betacam SP	8
CD-ROM	1
CD-RW	2
Compact cassette	12
DVD-R	41
DVD-Video	2
Slide	2
Sound disc - 78	1
Sound disc - Lacquer	5
Sound disc - LP	17
Umatic	294
VHS	179
Total	**690**

Conclusions:

- This collection is primarily Umatic, VHS, and film – a mix of training videos and medical history.
- While a significant number of these may be published works, their age and topic suggests that they would be sufficiently rare and would have research value from a medical history standpoint.
- The highest risk formats include:
  - Open reel video
  - Acetate-backed film
  - Lacquer discs
  - Umatic and VHS, dependent on age and condition
  - Optical media

Music

Format	Count
1 inch sound tape reel	5
1 inch video reel	5
1/2 inch sound tape reel	100
1/4 inch sound tape reel	9
1/4 inch sound tape reel - acetate	1,148
1/4 inch sound tape reel - polyester	1,967
16mm film reel	118
16mm film reel acetate	3
16mm film reel polyester	88

16mm full coat mag track	6
2 inch quadruplex video reel	16
35mm film reel	62
35mm film reel polyester	58
35mm full coat mag track	1
8-track cartridge	3
Audio-CD	12
Betacam	9
Betacam SP	11
Betamax	1
CD-R	207
Compact cassette	1,416
DAT	78
DVD-R	4
DVD-Video	1
MiniDV	7
Slide	8
Sound disc - 45	5
Sound disc - LP	21
Umatic	87
Umatic SP	1
VCR VC-60	1
VHS	141
Total	**5,599**

Conclusions:

- The overwhelming majority of the collection consists of open reel audio and compact cassettes, and the bulk of those come from four collections: Benny Goodman, Stanley & Helen Oakley Dance, Robert Shaw, and Howard Boatwright.

- There may be some crossover in collections with HSR here: All Music repository materials came from LSF and were labeled as belonging to Music, whereas on the shelves at the library there was less differentiation.

- A portion of the Benny Goodman collection has been digitized at a preservation level recently, though it has been included here for inventory purposes.

- The highest risk formats include:
    - Acetate-backed ¼ inch and ½ inch and 1 inch open reel audio
    - 1 inch and 2 inch open reel video
    - DAT
    - Umatic, dependent on age and condition
    - Optical media
    - Acetate-backed film

OHAM

Format	Count
1/4 inch sound tape reel	51
1/4 inch sound tape reel - acetate	12
1/4 inch sound tape reel - polyester	334
Audio-CD	1
Betacam	1
Betacam SP	131
Blu-Ray	1
CD-R	550
CD-RW	2
Compact cassette	1,194
DAT	193
Digital Betacam	2
DVCAM	1
DVD-R	12
Microcassette	11
MiniDisc	57
MiniDV	29
Hard Drive	1
Sound disc - LP	1
Sound disc - Other transcription	1
Test Pressing	11
Umatic	120
Umatic SP	1
VHS	139
XDCAM	1
Zip Disk	**2**
Total	**2,859**

Conclusions:

- The bulk of materials are compact cassette and ¼ inch audio reel oral histories. There is also a large group of CD-Rs that include some original documentation and research documentation on oral history subjects, and video oral histories captured on BetacamSP, Umatic, and VHS.

- There is very little commercial/published work in this collection. The majority is original recordings. Published or broadcast works (i.e., those not created by OHAM) exist for research or further information on subjects, and therefore may have an important role in the collection overall.

- Highest risk formats included:
    - DAT
    - Microcassette
    - MiniDisc
    - Umatic

- Optical media
- Acetate-backed ¼ inch open reel audio

HSR

Format	Count
1 inch sound tape reel	8
1/2 inch sound tape reel	9
1/4 inch sound tape reel	240
1/4 inch sound tape reel - acetate	1,325
1/4 inch sound tape reel - polyester	12,458
8-track cartridge	1
8mm film reel	1
Audio-CD	112
Audograph	54
Berliner	665
Betamax	1
CD-R	3,839
Compact cassette	3,971
DAT	80
Digital 8	2
DTRS (Hi8)	2,298
DVD-R	33
DVD-RW	198
Iomega REV Disk	55
Microcassette	7
Hard Drive	2
PCM-F1	13
S-VHS	43
Sound cylinder	440
Sound disc - 45	27
Sound disc - 78	419
Sound disc - Lacquer	2,709
Sound disc - LP	205
Sound disc - Other transcription	657
Sound wire reel	44
SoundScriber	146
Test Pressing	3,744
Umatic	2
VHS	120
Total	**33,928**

Conclusions:

- By far the largest collection inventoried in this phase, the bulk of materials are open reel audio, compact cassettes, CD-R, DTRS, lacquer discs, and assorted test pressing content.

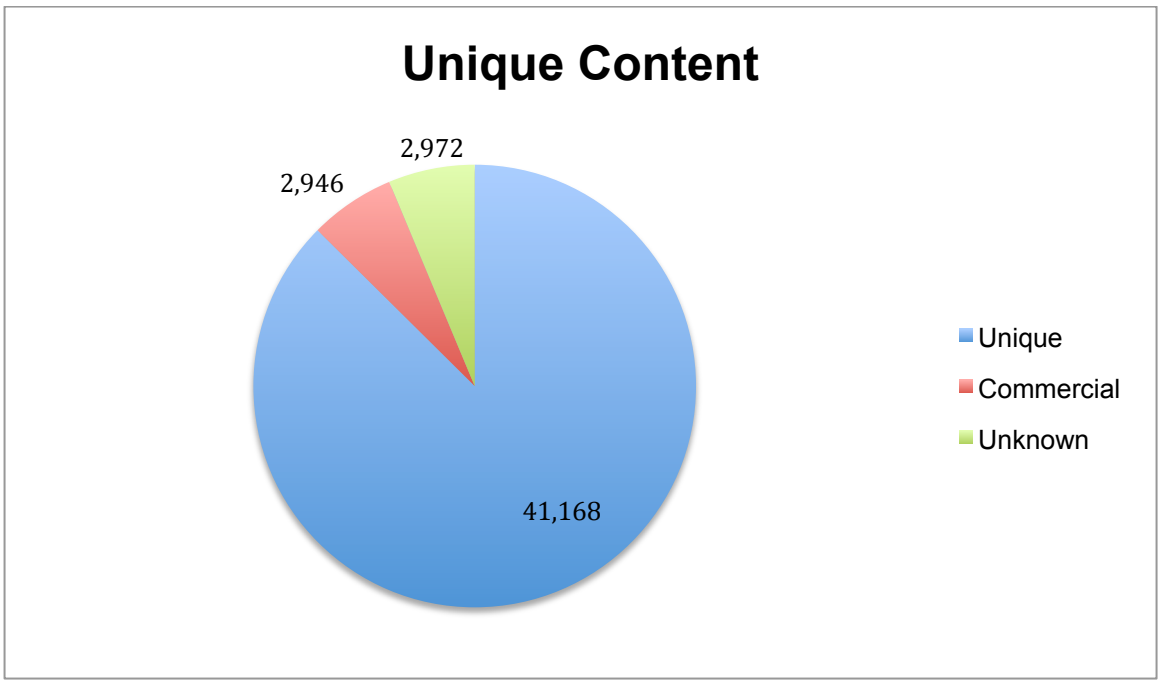
- As noted above, there are large portions of the open reel materials that are dubs of commercial 78s and lacquer discs, or even secondary dubs of other open reel items.

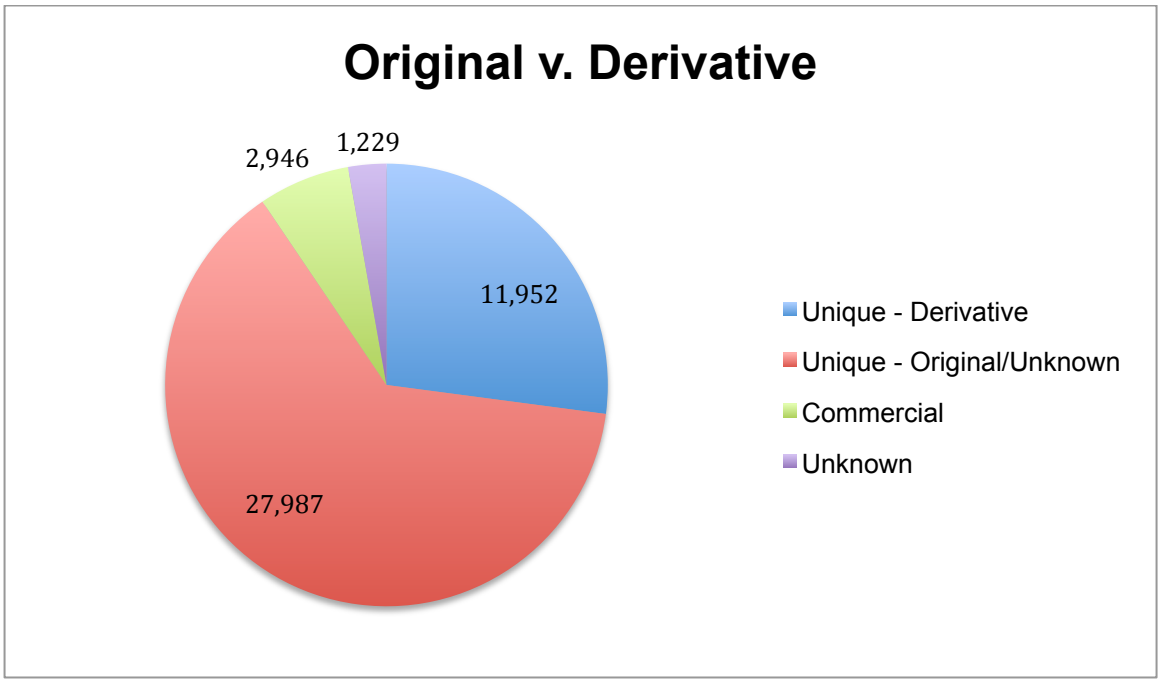
- While many of these and other distribution formats (cylinders, Berliners, the outcomes of test pressings) are published/commercial in nature, many are rare or help compose the breadth of a research collection.

- The highest risk formats include:
    - Acetate-backed open reel audio
    - Polyester-backed open reel audio, dependent on condition (the size of the collection increases the likelihood of tape with SBS or other degradation issues)
    - DAT
    - Optical Media
    - DTRS
    - Wire reel
    - Lacquer disc
    - Microcassette
    - PCM-F1
    - Cylinders (mitigated by previous transfers)

Unique Content

From initial analysis and best guess, 41,168 items (87.4%) would be considered unique or rare content, 2,946 (6.3%) commercial, and 2,972 (6.3%) unknown. In this inventory, “Unique” was used to refer to published materials that would be sufficiently rare to considered potential preservation efforts, depending on the intellectual importance of the content to the collection. However, among the “Unique” content, at least 11,952 (28.7%) were identified as “Derivative”, meaning they were dubs or transfers from originals. There may have been more, but a large percentage of items were unidentifiable as far as their original/derivative status. This would leave potentially 29,216 original items to address for prioritization. This number would be lowered when accounting for originals that have already been transferred or content that is not considered an intellectual priority.


Unique Content
2,972
2,946
41,168
Unique
Commercial
Unknown


Original v. Derivative
2,946
1,229
11,952
27,987
Unique - Derivative
Unique - Original/Unknown
Commercial
Unknown

Significant Preservation Risks

For audiovisual collections, preservation risks – here meant to encompass those risks that impact the condition of an item as well as the ability to access it both in the near-term and long-term – fall into three general categories: Degradation, Obsolescence, and Intellectual Control. In many cases these risk categories can be intertwined in a way that amplifies the overall risk. Ultimately the only long-term preservation strategy for audiovisual materials is reformatting. Approaches such as improving current storage conditions may help extend the physical longevity of an item, but that duration is limited and will eventually be eclipsed by obsolescence factors. Several of these factors are mentioned in the collection analysis above, but are reiterated and described in further detail here

Specific Risks: Degradation (& Physical Damage)

Audiovisual carriers are manufactured from a mix of organic and inorganic materials: plastics, metals, dyes, lubricants, and (often unidentified) chemical compounds. These materials decay at different rates and in different ways from one another, and in some cases can interact to instigate or accelerate overall decay.

Decay processes will manifest in several ways: brittleness, delamination of emulsion/binder from a base, color/image fading, chemical leeching, shrinkage, oxidization, and more. These processes can result in specific physical conditions (such as cupping, warping, breakage, oxide shedding, etc.), though the ultimate outcome is that there is a partial or total loss of the recorded image/signal, or the item becomes unplayable and is essentially lost.

Among the materials documented in the inventory, the most significant degradation risks include:

1. Acetate Transcription Discs: Though the discs in the HSR collection appeared to be in relatively good condition, the acetate lacquer coating these discs is very unstable and will eventually begin to flake and delaminate. Once this process begins the entire recording becomes essentially unplayable. Glass-based discs are also high-risk due to fragility.

2. Acetate-backed Audiotape: Acetate-backed tape shrinks and becomes brittle over time, leading to potential breakage during handling, delamination, and a decrease in the quality or loss of the audio signal.

3. Acetate-backed Film: Acetate-backed tape shrinks and becomes brittle over time, and is also prone to Vinegar Syndrome, leading to potential breakage during handling, delamination, image loss, color fading, and lack of playability.

4. Open Reel Video (1/2 Inch, 1 Inch, 2 Inch): As the oldest video format, open reel video has one of the highest risks of decay as the binder breaks down from SBS or other causes. 2 inch can be especially problematic because it is the oldest format, and 1/2 inch is especially prone to binder decay. 1 inch video was in use into the 1990s at least, so the impact of decay will be more dependent on age.

5. Videocassettes prior to ~1985 (Umatic, VHS, Betamax): There is no hard date for when magnetic media begins to break down, and there are variables such as chemical makeup and storage history that can delay or accelerate decay, but a general rule of thumb is that tape greater than 25-30 years of age is entering the high risk zone.

6. Audiocassettes prior to ~1980: Similar to videocassettes above, though audiocassettes have presented fewer issues than videocassettes. However, there is a preponderance of off-brand cassettes, specific “bad batch” brands, and models of tape that are at higher risk for binder decay.

7. Polyester-backed Open Reel Audiotape dating from the late 1970s – early 1980s: Polyester-backed tape will experience binder decay over time, but there was an especially bad period of manufacturing during this time period that has led to a greater degree of SBS in many brands and models.

8. Wire Recordings: Wire used for audio recording is very thin and can easily tangle or get damaged during handling and playback. Certain metal formulations are prone to oxidization.

Specific Risks: Obsolescence (& Other Technical Problems)

All audiovisual material is dependent on some kind of machinery for access. Some early technology is fairly simple and reproducible, but as time moved on formats became more proprietary and more dependent on complex machinery, and new formats were created, marketed, and discontinued more frequently. Once decks are no longer manufactured there is a period when surplus inventory and used items are still available. Over time those pools dry up, machines wear down, parts are not replaceable, and it becomes more difficult and more costly to access or transfer materials. Obsolescence is not a measureable date from the end of manufacture – much can depend on the market penetration of the format, the number decks manufactured, and the quality of the decks. Obsolescence can come slowly or it can happen suddenly.

Among the materials documented in the inventory, the most significant obsolescence risks include:

1. Open Reel Video (1/2 Inch, 1 Inch, 2 Inch): Open reel decks have not been manufactured for many decades in some cases. A small number of 2 Inch decks

remain, and 1/2 inch decks are becoming scarce as they break down and are scavenged for parts.

2. DAT: A short-lived, professional format. A relatively small number of decks were actually manufactured and sold, and those in existence are highly prone to failure.

3. DTRS: Similar to DAT, primarily used in production environments, thus limited in adoption and with limited equipment available. Digital formatting of the audio signal also causes potential issues. Other digital audio formats such as PCM-F1 (Betamax), ADAT (VHS), and 1610/1630 (Umatic) can be grouped here as well.

4. Wire Recording: Wire recordings were mostly in use during the early-mid 20th century -- working playback machines are difficult to find and experience in handling the format is limited.

5. MiniDisc: A prosumer format with a short period of popularity. Proprietary playback machines and encoding format can cause transfer issues.

6. Lacquer Discs: While turntables are not necessarily in short supply and the technology is not as difficult to recreate, there was no standard needle size used for transcription discs, and the proper size is usually not recorded on the item. Difficulty in identifying the correct needle and the risks involved in playing a lacquer disc more than once may impact the quality of transfers.

7. D2: Short-lived professional format with limited adoption. Electronic complexity of decks means they are prone to failure.

8. MII: Short-lived professional format with limited adoption. Few decks in existence.

9. Data Storage Formats (Iomega REV, Zip Disk, Floppy Disk, Etc.): Hardware is no longer being manufactured, associated software may be difficult to find, and there may be OS conflicts with legacy formats.

10. Trending Obsolescence -- Umatic, MiniDV, Betamax, Hi8, Digital8: These were fairly widespread professional and prosumer formats with a large number of decks available for playback. However, as digitization increases over the coming years fewer decks will be available on the resale market and more and more will break down, leading to significant near term concern regarding longevity.

Specific Risks: Intellectual Control

Lack of intellectual control for audiovisual collections is a preservation risk. Items that are not findable, identifiable, or accessible due to lack of annotations or lack of playback equipment will not be requested and therefore will not be considered a prioritization for preservation. Additionally, preservation efforts cannot be budgeted and planned without the basic capability to perform selection/prioritization or to communicate technical information about a collection (formats, number of items, durations, condition issues) to vendors or Preservation departments.

The inventory generated in this project is the first step to getting the intellectual control that will enable preservation decisions and planning. However, there are still some decision points that will need to be evaluated:

1. Intellectual valuation for prioritization.

2. Determination of whether past reformatting efforts are currently sufficient or if reformatting should be re-performed.

3. If multiple copies exist, which should be the source for reformatting (Best Copy Selection).

4. If a particular work is represented by production materials, a final master, and/or a distribution version, which should be considered the “copy of record”?

Prioritization Schema

General Guidelines:

- Prioritize unique/rare materials over commercial, unless there are cases where a commercial/published work is critical to the use/understanding of the collection or is sufficiently rare that it would be considered a high preservation priority.

- Prioritize content that has not been reformatted previously or that has not been reformatted by other institutions. Cases where previous reformatting occurred for access purposes only may be evaluated separately based on the intellectual content and preservation risk of the original source item.

- Intellectual prioritization may take precedence over technical factors, however there are certain formats that will have very short-term longevity and that should be intellectually appraised and addressed sooner than others.

- Collection policy and funding opportunities will have an impact on selection that may supersede technical factor or may support fundraising attempts.

- Reformatting for preservation results in a different quality than reformatting for access only. Access only copies can become the ultimate preservation master if no other options are available, but that should not be the de facto approach.

Technical Prioritization:

Audio

Format	Note
Sound disc - Lacquer	Highly unstable and easily damaged format.
DAT	Limited number of decks and frequent playback issues.
DTRS (Hi8)	Limited number of decks and frequent playback issues.
PCM-F1	Limited number of decks and frequent playback issues.
MiniDisc	Proprietary encoding and hardware accessibility issues.
Sound cylinder	Fragile medium, though there have been many international projects to digitize commercial examples.
1/4 inch sound tape reel - acetate	Severe decay and damage issues with acetate backing.
1 inch sound tape reel	Unidentified backing. May include acetate.
1/2 inch sound tape reel	Unidentified backing. May include acetate.
1/4 inch sound tape reel	Unidentified backing. May include acetate.
1/4 inch sound tape reel - paper	More stable than acetate, but represents the oldest magnetic media.
1/4 inch sound tape reel - polyester	Binder stability issues dependent on age and storage conditions
Sound wire reel	Fairly stable physically, but difficult to transfer.
Microcassette	Small, thinner tapes are more prone to damage or poor audio reproduction during playback.
Compact cassette	Age, condition, and content dependent. May include numerous commercial items.
CD-R	Increasing stability and encoding issues, however many of these may represent access copies for previously digitized works.
CD-RW	Increasing stability and encoding issues, however many of these may represent access copies for previously digitized works.
Sound disc - Other transcription	May include lacquer or metal discs.
SoundScriber	Soft vinyl is stable but limits amount of playback.
Audograph	Soft vinyl is stable but limits amount of playback.
Test Pressing	Requires significant intellectual valuation.
Audio-CD	Commercial content
Berliner	Commercial content – dependent on intellectual value or rarity
Sound disc - 78	Commercial content – dependent on intellectual value or rarity
Sound disc - 45	Commercial content

Sound disc - LP	Commercial content
8-track cartridge	Commercial content

Video

Format	Note
2 inch quadruplex video reel	Severe obsolescence and decay issues.
1/2 inch video reel	Severe obsolescence and decay issues.
VCR VC-60	Severe obsolescence and decay issues.
D2	Severe obsolescence and technology issues.
MII	Severe obsolescence and technology issues.
Umatic	Condition dependent on age, brand, and storage conditions. Availability of hardware will be a growing concern.
Umatic SP	Condition dependent on age, brand, and storage conditions. Availability of hardware will be a growing concern.
1 inch video reel	Condition dependent on age, brand, and storage conditions. Availability of hardware will be a growing concern.
Digital 8	Obsolescence and low quality transfers are a concern.
Hi8	Obsolescence and low quality transfers are a concern.
Betamax	Heightened obsolescence concern, and well as condition due to age.
MiniDV	Low quality transfers are a concern.
DVCAM	Low quality transfers are a concern.
VHS	Condition dependent on age, brand, and storage conditions. Transfer can be poor due to low quality of format and hardware.
VHS-C	Condition dependent on age, brand, and storage conditions. Transfer can be poor due to low quality of format and hardware.
S-VHS	Condition dependent on age, brand, and storage conditions. Transfer can be poor due to low quality of format and hardware.
CD-ROM	Increasing stability and encoding issues, as well as technological obsolescence.
DVD-RW	Increasing stability and encoding issues, however many of these may represent access copies for previously digitized works. Rewritable discs are higher decay risk than –R discs.
DVD-R	Increasing stability and encoding issues arising, however many of these may represent access copies for previously digitized works.
Betacam	Stable at this point in comparison to other formats, but binder decay will begin to present and is already a problem in certain tape formulations.
Betacam SP	Stable at this point in comparison to other formats, but binder decay will begin to present.
Digital Betacam	Stable at this point in comparison to other formats, but

	binder decay will begin to present and technology issues are a rising concern.
Blu-Ray	Still relatively new format. Risks not yet identified.
XDCAM	Still relatively new format. Risks not yet identified.

Film

Format	Note
Acetate backed color film	Acetate is a much greater risk than polyester, and color film fades and has image loss issues that are not as severe as with B&W. However, the age of a B&W film may lead to greater level of acetate decay.
Acetate backed B&W film	
Polyester backed color film	
Polyester backed B&W film	
Full coat mag track	This is actually a high-risk format, but the transfer is dependent on many other factors, such as best copy selection and the relationship to the image element. The audio would likely not be transferred without also doing the image, so they would be selected in tandem.

Other

Format	Note
Zip Disk	All digital formats should be moved to a file-based environment as soon as possible. Formats that require specialized drives such as Zip Disks, REV Disks, and Floppy Discs, may be more of a priority depending on the digital formats they contain.
Iomega REV Disk	
Floppy disc	
Thumbdrive	
External Hard Drive	
Film Strip	Both will likely suffer from color fading and potential shrinkage if they are acetate. Content review and collecting policy will be important factors.
Slide	

Digitization & File Storage Needs

Digitization Specifications

Target Format Recommendations

For the past several decades, archival practices include five objects as part of a digital preservation strategy:

Object	Purpose
Source Asset (analog or digital object)	The original or highest quality existing copy of an audiovisual asset from which the signal is digitized or migrated. After digitization, these are prepped for long-term storage and infrequently accessed. As part of a preservation strategy, they are typically maintained in long-term storage conditions as a backup option.
Preservation Master (digital file)	High-resolution file created from the digitization of the Source Asset. The Preservation Master is used to generate the Access Master. The Preservation Master takes the place of the Source Asset in future reformatting/migration unless the file is lost or corrupted or it is determined that a new Preservation Master must be created. Preservation Masters are infrequently[1] accessed except for regeneration of Access Masters and for periodic bit preservation analysis. As such, offline or data tape storage can be considered.
Access Master (optional digital file)	Also referred to as the Mezzanine Copy. The Access Master is a medium-sized file, small enough to be easily handled by most computer and editing systems but of a high enough resolution to be a quality source for creating Access Copies and for editing into new video and digital content. Can be stored online or nearline depending on frequency and method of access. May not be necessary to generate if Preservation Master level files are accessible and a transcoding workflow is in place.
Access Copy (digital file)	Low resolution file that can be distributed over the internet or maintained on some other media used for frequent or easily obtainable listening/viewing. The Access Copy can be regenerated from the Access Master and does not require a high level of preservation.
Preservation Master (digital backup)	At least one backup copy of the Preservation Master should be stored in a geographically separate location from the primary Preservation Master.

A Note on Access Masters

Access Masters are an optional middle tier; while not considered adequate to serve the purpose of preservation, they are of a high enough quality to satisfy most production needs, but small enough to allow easy manipulation. They are practical when the Preservation Master files are difficult to work with or are stored offline and require a lot of time and/or effort to restore. The inclusion of Access Master files in a digitization plan adds to the plan's cost; the amount of required digital storage increases and some vendors charge more for multiple derivatives.

[1] On the scale of years, not multiple times per year for the same file.

As YUL does not currently see production as a potential part of their workflow, we do not recommend that they create Access Master files for their video and audio assets. Access Master files do, however, make sense for film assets, as their Preservation Master files are large, complex, and difficult to manipulate.

Recommended Target File Formats for Digitization (Summary)

Video	
Preservation Master	
DVCAM, MiniDV	AVI wrapper; Native DV codec; Native audio Migrate directly with no transcoding Maintain the original ancillary data within the essence such as original time code and closed captioning
DVD	Migrate directly with no transcoding
High Value SD Content	AVI wrapper; 10-bit FFV1 version 3 codec Uncompressed PCM audio tracks, 48kHz Maintain the original aspect ratio, recording standard, interlacing, number of audio channels, and ancillary data within the essence such as original time code and closed captioning
Reference SD Content	QuickTime DV codec 25Mbps Maintain the original aspect ratio, recording standard, interlacing, number of audio channels, and ancillary data within the essence such as original time code and closed captioning
Access Master	
Not Recommended	
Access Copy	
All Formats	MPEG4 wrapper; H.264 5.0Mbps Uncompressed AAC audio, 44.1kHz, 256kbps Maintain the original aspect ratio, recording standard, frame rate, and number of audio channels. File should be optimized for streaming and de-interlaced. Any existing time code and closed caption data should be mapped to the time code track and timed-text tracks within the MPEG-4 wrapper.
Audio	
Preservation Master	
CD	Broadcast Wave File wrapper Migrate directly with no transcoding
DAT	Broadcast Wave File wrapper Maintain original sample rate and bit depth Migrate directly with no transcoding
All Other	Broadcast Wave File wrapper; 96kHz, 24-bit, uncompressed PCM
Access Master	
Not Recommended	
Access Copy	
All Formats	MP3 wrapper; 44.1 kHz; 256kbps

Film	
Preservation Master	
8mm/16mm High Value Silent & Image with Sound	DPX wrapped 2k Uncompressed RGB 4:4:4 Uncompressed PCM audio tracks (if applicable)
35mm High Value Silent & Image with Sound	DPX wrapped 4k Uncompressed RGB 4:4:4 Uncompressed PCM audio tracks (if applicable)
Reference Value Silent & Image with Sound	AVC Intra 100, RGB 4:2:2, 23.98 FPS Progressive with pillars for SD content to retain aspect ratios Uncompressed PCM audio tracks (if applicable)
Access Master	
High Value Silent & Image with Sound	AVC Intra 100, RGB 4:2:2, 23.98 FPS Progressive with pillars for SD content to retain aspect ratios Uncompressed PCM audio tracks (if applicable)
Access Copy	
Silent & Image with Sound	MPEG4 wrapper; H.264 5.0Mbps Uncompressed AAC audio, 44.1kHz, 256kbps

The recommended Video Preservation Master file format for cases where the content is unique or of high institutional value is 10-bit FFV1. FFV1 may also be used for distributed works or objects where YUL does not hold the master, but DV25 may be adequate in these cases where access is the primary purpose of reformatting rather than long-term preservation.

It is a similar case with film scanning where DPX 2k or 4k uncompressed is recommended for materials of high institutional value, and AVC-Intra 100 is recommended for published works or works not considered suitable for long term preservation. AVC Intra 100 offers a high quality image that some consider adequate for any Preservation Master, but its adoption is not as widespread and it does not offer a 4:4:4 color or 4:3 aspect ratio. If 2k/4k uncompressed is used, the generation of an Access Master is recommended due to the large size and complexity of 2k and 4k DPX files. AVC Intra is also somewhat specialized, and while it is not readily accessible using standard default players, it is widely used within video production and therefore accessible using video editing software and plug-ins. It is assumed that the likely use-case for the Access Master level file in this instance will be for production purposes and will therefore work well.

In Appendix A, YUL File Size Calculator, there are options for viewing resulting file sizes for Uncompressed, JP2K, FFV1, ProRes 422, DV25, MPEG-2, and AVC Intra video files, 96kHz and 48kHz sample rates for audio, and 2k uncompressed, 4k uncompressed, 2k compressed, 4k compressed, AVC-Intra 100, and IMX formats for scanned film. Each will have its pluses and minuses as far as capabilities, interoperability, file size, and future migration. Ultimately, the decision will be informed by the infrastructure and resources available to YUL.

Video

Preservation Master

Recommended: AVI wrapper; FFV1 lossless codec

Alternate 1: QuickTime .mov wrapper; 10-bit YUV 4:2:2 uncompressed v210 codec

Alternate 2: QuickTime DV Codec 25Mbps

Why FFV1: Open standard that uses lossless compression to shrink file sizes without loss of image quality. Considered by some to be more useable than the other lossless compression format, JPEG2000. Adoption by archives and cultural heritage organizations is spreading. User support is growing.

Why Not FFV1: Common wrapper formats (AVI and Matroska) do not support time code and other ancillary data. Technically complex standard and more specialized software needed.

Why Uncompressed: Current de facto standard for video preservation. Widely adopted and accessible by most software. Presents high quality and minimal generation loss in reformatting. Leaves future migrations more flexible and with potential for higher quality by not locking into a compression scheme.

Why Not Uncompressed: Results in very large file sizes (~105GB per hour of content) that may result in the need to store petabytes of data.

Why DV25: Widely adopted and accessible by most software. Simple technical standard. Vastly smaller file sizes than uncompressed or even other compressed formats.

Why Not DV25: Compressed format. Generation loss. Fewer options for future migration.

For DV Family and Optical Disc

Recommended: AVI wrapper with native codec, migrated directly via Firewire or migrated directly with no transcoding for optical discs.

Alternate: None

Why Native Codec/Direct Migration: Formats such as DVCAM and miniDV are already digital and the codecs used are already compressed. The codecs are also in wide-use and very accessible using a wide-ranging and readily available toolset.

Transcoding to uncompressed will add nothing to the quality of the image, and digitizing as with other video formats will result in the loss of metadata contained in the bitstream.

Why Not Native Codec/Direct Migration: N/A

Access Master

Not recommended as part of the preservation strategy. There is no compelling use case for this file since production is not a frequent activity at YUL.

Access Copy

Recommended: MPEG4 wrapped H.264 at 5.0Mbps

Why MPEG4/H.264: Pervasive and widely adopted. High quality, flexible standard. Web distribution ready. 5.0Mbps is the current standard used by many online distribution platforms. Lower bitrate may result in a sooner need to create newer access copies that meet expected web standards.

Why Not MPEG4: N/A

Audio

Preservation Master

Recommended: Uncompressed PCM Broadcast Wave File (BWF) at 96/24

Alternate: Uncompressed Broadcast Wave File (BWF) at 48/24

Why 96/24: Widely adopted and accessible by most software. Presents high quality and minimal generation loss in reformatting. Leaves future migrations more flexible and with the potential for higher quality by not locking into a compression scheme. Highest bit rate and sampling rate in typical use. Audio preservation industry standard.[2]

Why Not 96/24: Results in comparatively larger file sizes. Depending on the source content there may be little discernable quality to the ear between 96kHz and 48kHz.

Why 48/24: Still offers the accessibility and flexibility, but at a smaller file size.

[2] http://www.iasa-web.org/tc04/audio-preservation

Why Not 48/24: File size savings are minimal. It limits future use and options more so than 96/24. Depending on content, difference is likely to be notable.

For DAT and Optical Disc

Recommended: Uncompressed PCM Broadcast Wave File (BWF) at native sample rate and bit depth, migrated directly via AES or migrated directly with no transcoding for optical discs.

Alternate: None

Why Direct Migration: Formats such as DATs are already digital and the sample rate and bit-depth were defined at the point of creation. Transcoding to a higher sample rate or bit-depth would add nothing to the quality of the audio.

Why Not Direct Migration: N/A

Access Master

Not recommended as part of the preservation strategy. There is no compelling use case for this file as production is not a frequent activity at the YUL.

Access Copy

Recommended: 256kbps MP3

Why MP3: Pervasive and widely recognized adopted standard for delivery of audio.

Why Not MP3: N/A

Film

Preservation Master

Recommended: DPX wrapped 2k (8mm/16mm) or 4k (35mm) Uncompressed RGB 4:4:4 10-bit log with Broadcast Wave File wrapped uncompressed audio

Alternate 1: AVC Intra 100, RGB 4:2:2, 23.98 FPS Progressive with pillars for SD content to retain aspect ratios

Alternate 2: IMX at 50Mbps, progressive and 25fps (MXF wrapped MPEG2 422P@ML, I-Frame)

Why Uncompressed: Widely adopted in the film production and post-production domain. Presents high quality and minimal generation loss in reformatting. Leaves future migrations more flexible and with potential for higher quality by not locking into a compression scheme. Most sustainable format for capturing at preservation-level resolution.

Why Not Uncompressed: Results in very large file sizes (1TB per hour of content for 2k) that may result in the need to store petabytes of data. Technically complex and not as widely usable across systems. Complexity plus size makes it less practical to use as a source for on-the-fly transcoding.

Why AVC-Intra 100: Smaller file sizes than uncompressed. High comparable quality to other compressed formats.

Why Not AVC-Intra 100: Not as widely adopted and may not be a long-term solution. Issues with handling complexity of MXF. Does not completely reflect the color depth (422 vs 444) or frame rate (23.98 FPS vs 24 FPS) of film. Lossy format.

Why IMX: High quality image production adopted by broadcast. Lower file size. Easier to access than AVC-Intra 100.

Why Not IMX: Quality and options available do not maintain integrity as well as other options. Using a lossier compression than AVC-Intra 100, IMX leads to more loss over time as it is forced to migrate to new formats.

Access Master

Only recommended if using 2k/4k uncompressed as a Preservation Master. 2k and 4k files are large and complex, and the strategy to use the Preservation Master as the source for generating Access Copies is not as viable.

Recommended: AVC-Intra 100, RGB 4:2:0, 23.98 FPS Progressive

Alternate: No alternative

Why AVC-Intra 100: Smaller file sizes than uncompressed. High comparable quality to other compressed formats. Production quality.

Why Not AVC-Intra 100: N/A

Access Copy

Recommended: MPEG4 wrapped H.264 at 5.0Mbps

Why MPEG4/H.264: Pervasive and widely adopted. High quality, flexible standard. Web distribution ready. 5.0Mbps is current standard used by many online distribution platforms. Lower bitrate may result in a sooner need to create newer access copies that meet expected web standards.

Why Not MPEG4/H.264: N/A

File Sizes Per Hour of Content

Video	
Preservation Master	
QT .mov Uncompressed SD	105 GB
FFV1 Lossless Compression	36 GB
25Mbps DV	14GB
Native Codecs	Most are around 11-12GB
Access Copy	
MPEG4 H.264 5.0Mbps	2.25 GB
Audio	
Preservation Master	
96/24 BWF	2.07 GB Stereo / 1.04 GB Mono
48/24 BWF	1.04 GB Stereo / .52 GB Mono
DAT Native Codec	0.65 GB Stereo / .33 GB Mono
Access Copy	
256kpbs MP3	0.12 GB
Film (Image & Image with Sound)	
Preservation Master	
2k Uncompressed	1 TB
4k Uncompressed	4 TB
AVC Intra 100	56.6 GB
MPEG-2 50Mbps I-frame	94 GB
Access Master	
AVC Intra 100	56.6 GB
Access Copy	
MPEG4 H.264 5.0Mbps	2.25 GB

Using the File Size Calculator

A file size calculator can be found in Appendix A, YUL File Size Calculator. This spreadsheet contains all audiovisual assets imaged during this project, meaning potential duplicates or non-preservation worthy materials have not been removed from the counts. Slides, film strips, and other non-audiovisual items considered outside of the scope of the project were also not included. Durations and film lengths were based on notations on the Source Assets or, in their absence, standard media capacity. Durations

based on capacity reflect what would be the maximum capacity of the media object, which in many cases will be more than the duration of the actual content.

The spreadsheet is organized by library rather than format. Each of the six libraries included in this project – Arts, Historical Sound Recordings, Medical, Music, Oral History of American Music, and Divinity – have their collections on a separate sheet. There is then a "Totals" sheet, which combines all collections, giving an overview of all YUL's audiovisual holdings.

The file calculator estimates the total amount of digital storage required for digitized assets, assuming a high resolution Preservation Master with one backup copy and a lower resolution Access Copy that can be used for streaming and access (film assets also receive an Access Master file). At the bottom of the "Totals" sheet is a chart where it is possible to change what the target preservation format would be in order to see the storage impact of selecting lower resolution or compressed formats. These are typically not recommended except in cases where the Source Asset is very low quality or not a high priority, or where an informed decision has been made about accepting the limitations or risks of such formats.

Next Steps

- Intellectual prioritization by repositories
    - The outcome of this inventory will assist the Preservation Department in prioritizing materials at the technical level. However, the intellectual component will be an important part of selection. The size and contents of collections, as well as the timeline and cost for reformatting (and the cost of digital storage/preservation moving ahead), will likely mean that not everything can or should be reformatted. If we are looking at a 5-15 year window for audiovisual reformatting as most experts agree at this time, the repositories will need to undergo an intellectual assessment of their collections in the very near-term in order to have the best opportunity at saving the greatest amount of materials.

    - Unlike technical prioritization, intellectual prioritization cannot be performed at the highest level. The six repositories included in this inventory project – Arts, Medical, Divinity, Music, HSR, and OHAM – must each prioritize their own collections based on a combination of content value and use value. Then, with the assistance of the Preservation Department, the prioritization may be further honed using technical considerations based on format, age, and condition.

- Determine policy approach to commercial works and to works that had been previously transferred to tape or optical disc

    - The items inventoried contained commercial works, as well as items that had already been transferred to another storage format (open reel tape, CD, DVD, etc.). Either as a working group or at the individual repository level, a decision will need to be made on the priority of reformatting commercial works and previously reformatted works, as well as what should be the source of works to be reformatted (e.g., the HSR lacquer disc collection that had previously been transferred to 1/4 inch open reel audio and the question of should digitization be performed using the original lacquer disc or the subsequent open reel?).

    - There will be differing opinion among the repositories on this given the formats and collection scope they maintain. For example, HSR has a very large commercial 78 RPM disc collection, which fits into their collection policy to server for researchers, whereas Divinity's collections are more likely to be original recordings or limited run production materials that represent an institution's output and therefore may be important for researching an entity. In the case of all collections there may be commercial works included in a named collection, in which case the determination must be made of whether the actual content is of importance to the collection, or just the physical item and record of its existence (especially in instances where that content can be accessed through other means).

- Address housing in B20B
    - A majority of the boxes stored in B20B make up the loosely defined HSR collection known as Test Pressings. The boxes in use are quite large, probably originally chosen due to their ability to hold 16" discs. However, due to their size, almost all of them are overfilled. The weight of the discs is much too great for the boxes and for staff members to lift down from shelves. Handles were ripped and the bottoms of boxes sagged. If these materials are considered to be of value, they should be intellectually arranged and rehoused into smaller boxes that can be physically handled without endangering the contents for the staff members handling them.

- Determine digitization specifications
    - Recommendations for target specifications are provided above. These are only recommendations. If YUL maintains current specifications or specifications must be designed to integrate with any existing workflows/systems, those requirements would necessarily take precedence until any institutional review is performed. A dialog should be begun with YUL's IT department to determine the necessary specifications.

- Begin near-term prioritization to begin immediate fundraising/planning
    - There is no doubt that digitization, followed by digital storage/preservation will require significant budget allocations. Along with the need for

intellectual prioritization in the very near-term, fundraising and planning will need to begin soon in order to set the preservation process in motion in a timely manner. Oftentimes, it can prove beneficial to select a popular collection to start the fundraising process, as outside donors tend to give in accordance with their interests. YUL may decide to prioritize in this fashion in the near-term. Regardless of approach, interest must be generated and funds raised as soon as possible.