Assessment of the APCOB Collection

A Sub Collection of NYU's Special Collection of Indigenous Media

By Caroline Roll
Digital Preservation
Cine-GT 1807
Nicole Martin

Introduction

This assessment evaluates the APCOB (Apoyo Para el Campesino-Indígena del Oriente Boliviano) Collection, a newly acquired sub-collection of the NYU Special Collection of Indigenous Media at Elmer Holmes Bobst Library. Bobst is the main library in New York University's extensive library system, and there are a variety of special collections in its holdings. The Special Collection of Indigenous Media was created upon receipt of a large collection from the Smithsonian National Museum of the American Indian in 2015. It consists mostly of works created by indigenous filmmakers for the museum's Native American Film and Video Festival (1979-2011), along with a number of other audio and visual works donated to the museum. NYU is in the process of obtaining permissions from filmmakers to provide access to these works for research purposes, as well as to reformat and preserve them. It is important treat this material with cultural sensitivity, and library staff members are considering how to include the filmmakers and relevant peoples in describing this content. There is also interest in creating a union catalogue of the indigenous works held across various institutions.

APCOB was founded in 1980 to support indigenous peoples of Bolivia. They describe themselves as a "non-governmental humanist organization" focused on intercultural education, agricultural sustainability, and the assertion of indigenous rights in political matters.[1] The collective has produced books, documentary films, and interactive multimedia over the years. Their audiovisual content was sought out and obtained by Angela Carreño, Head of Collection Development and Librarian for Latin American & Caribbean Studies, along with the permission for NYU Libraries to preserve and provide access to the material.

[1] "Apoyo Para El Campesino-Indigena Del Oriente Boliviano." APCOB. Web.

The following assessment provides an overview of the collection's content and condition, an explanation of how it has been handled so far, and recommendations for how to care for it going forward. This work was conducted in November and December 2016.

Content of the Collection

There are a variety of works in this collection. Many of them are documentaries highlighting specific indigenous groups, like the Chiquitano, Isoseño-Guaranies, and Ayoreo peoples. Others document educational efforts, such as a series promoting literacy in indigenous and Castellan languages. Many of these works are in Spanish, but there are also some in English and German, as well as in several indigenous languages. Many were produced in partnership with organizations such as UNICEF Bolivia, Cordaid, Bolivia's Ministry of Education, the Confederation of Indigenous Peoples of Bolivia (CIDOB), the U.S. Agency for International Development (USAID), and the Government of Finland. In addition to the video works, there are also interactive CDs and DVDs intended for educational purposes.

Physical Appraisal

This collection consists entirely of digital optical disc media. There are a total of 124 items, most of which are DVDs, and the rest are interactive DVDs and CDs (DVD-I and CD-I). Disc media is not archival quality, and it is known to have a relatively short lifespan. DVDs and CDs both have a standard diameter of 12cm, and both contain a reflective metal layer and a polycarbonate substrate layer. Risk of metal layer corrosion is common, and other signs of deterioration include pinholes in the metal layer, thinning, and discoloration. In addition, a DVD

is actually two discs bonded together, and some worry that the bonding layer could weaken or deteriorate, though this hasn't been observed in test situations.[2]

There is limited empirical data on disc longevity, and much of this information comes from manufacturers. Some have claimed that their discs will last 100 to 200 years, likely a gross exaggeration, though longevity can differ depending on the manufacturing brand.[3] In reality, the Library of Congress estimates that disc media can be expected to last about 30 years in favorable storage conditions of 25ºC and 50% relative humidity. George Blood Audio Video Film suggests an average lifespan of only five to ten years, which can be extended with proper storage conditions.[4] Disc media is also dependent on having both hardware and software to read its contents.

It is unknown how old the APCOB discs are and how they were stored before their transfer to NYU. The collection was sent to NYU in a bag, and half of the discs were stacked together, while the other half were wrapped in paper jackets. A physical inspection of these discs shows no significant damage, but many do have obvious scratches. It is unclear if these are a result of the transfer process or not. Many of them were tested in a computer disc drive and played normally, so the physical damage currently appears to be minor. However, one disc, titled *El Mundo de los Quechuas, Jalq'a y Tarabuco* (NYU2152_01), was discovered to be blank when inserted into the drive. It will be important to make sure there are no other similar anomalies in the collection.

Some of these discs have content that was created on older physical media. Simply from looking at the video images, many of them appear to have been transferred from videotape

---

[2] Morel M. Preserving Write-Once DVDs: Producing Disk images, Extracting Content, and Addressing Flaws and Errors. *George Blood Audio Video Film*. 2014.
[3] Iraci, Joe. 2005. "The Relative Stabilities of Optical Disc Formats." *Restaurateur* 26 (2): 134-150.
[4] Morel M. Preserving Write-Once DVDs: Producing Disk images, Extracting Content, and Addressing Flaws and Errors. *George Blood Audio Video Film*. 2014.

formats. Transferring content to DVD generally results in loss of quality, but it is unknown if the originals versions still exist. Regardless, NYU has been tasked with caring for these discs, and no better copies are available to them. The only course of action is to protect the physical media and create digital files of the content within.

Collections Care

It is unknown how these discs were stored before their transfer to NYU, but they were not in proper housing when they arrived. During the course of this assessment, each item was inventoried and labeled, and they were individually rehoused in new DVD cases. These will be stored in boxes and taken to an offsite, climate-controlled storage location where the Special Collection of Indigenous Media is held.

Due to the questionable longevity of disc media, it is considered good practice to create digital files to better ensure the preservation of their contents. One method is to convert it to a digital video file format, which can be streamed online or viewed in a player application. There are many different formats to choose from, and it is important to also consider the file codec. A codec is used to compresses a file so that it takes up less storage space. Some codecs offer lossless compression, thus maintaining the complete data in the file. Others are lossy but can still be acceptable for access copies. However, this conversion only extracts the video content, and does not include other features that may be on the DVD, like a menu or chapters. A good option for producing master files is the ISO file format. The process of converting to ISO is sometimes referred to as cloning because it creates an exact copy of the data stored on the disc.[5] This option is great for DVDs with a simple structure, as the DVDs in this collection are. However, interactive DVDs and CDs like the ones in this collection are much more complex, and there is

[5] Morel M. Preserving Write-Once DVDs: Producing Disk images, Extracting Content, and Addressing Flaws and Errors. *George Blood Audio Video Film*. 2014.

not a clear solution for creating digital files of their content. This process and its challenges are detailed further in the section “Creating Digital Files.”

Arrangement, Intellectual Control, and Metadata

Items in the Special Collection of Indigenous Media have been given unique identifiers following the template NYU0001_01. The number after the underscore indicates the number of copies of that item in the collection. It was decided that the APCOB items would have IDs that follow this pattern as well. In addition, digital versions of these items follow these same naming conventions - NYU0001_01.iso and NYU0001_01.mp4 - to ensure that the file can be easily associated with its physical version.

Physical Control

The physical items in the collection will be stored in a secure location offsite. Specific boxes may be requested at any time and are usually delivered to the library on the next business day. The library is in the process of establishing a barcode identification system for the Special Collection of Indigenous Media because there has been confusion in the past when requesting boxes from storage. Some boxes have duplicate ID numbers, and some have been hand-written or crossed out, so staff members often have to guess which box is the correct one to send. A barcode system will streamline box identification and enable storage staff to work more efficiently.

Currently, digital files related to the collection are stored on Google Drive and NYU Box. The working inventory is stored on Google Drive, while NYU Box is used for periodic backups. Test disc image files have been stored in both systems, but further consideration will need to be

given for long-term storage of digital master and access files. Google Drive only offers 15GB of free storage, so it is probably best to only use the service for working documents related to the collection. NYU Box allows unlimited storage with a maximum file size of 15GB, and this is being considered a "staging area" for these files while the library considers how best to approach file storage and fixity. Angela Carreño plans to consult with the newly created Archival Collections Management (ACM) department within Knowledge, Access, and Research Management Services (KARMS). ACM intends to provide archival technical services to the university's archives and special collections in a unifying approach. They will help implement standards for archival processing, descriptions, and collections management throughout the library, which will be incredibly helpful for Angela and her coworkers as they move forward with the Special Collection of Indigenous Media.

Rights Status

The APCOB Collection was purchased by NYU with the understanding that the library would need permissions for preservation work and for allowing access to researchers. Elizabeth Weatherford of the National Museum of the American Indian has been working closely with NYU during the transfer of the indigenous media collection, and has also taken part in the ongoing dialogue with APCOB. Emails between her and the Director of APCOB, Jorge Riester, indicate that Riester understands NYU's goals with the collection, that is, to digitize the works and provide access to them within the university. He signed a permission form for one title in the collection, but permissions still need to be secured for the rest of the collection. The challenge is that APCOB is a collective, and most of these films don't give credit to a single filmmaker, only crediting APCOB as the producer. Thus, the copyright status of many of these works is

uncertain, and the library is still in the process of working out this problem. It is important to note that, while test preservation files of some DVDs were made during the completion of this assessment, further preservation work should not continue until proper permissions documentation is received.

Creating Digital Files

During the course of this assessment, time was allotted to test out different options for digital file creation. There was not enough time to create master files for every disc in the collection, so my intent is to leave the library with options that a new intern can work with.

When extracting content from optical disc media, one can use tools in the command-line interface (CLI) or by using a graphical user interface (GUI) application. These options vary depending on the operating system used. The workstation used in this assessment is a Dell computer running Windows 7, and because I am only familiar with the command-line in Mac computers, I chose to focus on GUI applications for these tests.

There are a number of free applications to choose from. I settled on a free and open-source application called Infrarecorder, which is licensed under the GNU General Public License Version 3.[6] It quickly and easily creates ISO files, which can be considered the master copies of items in this collection. I created several disc images of DVDs from the collection, and found that this was about a two to five-minute process depending on the amount of content on each disc. ISO files can be viewed in DVD player software, and VLC player is a common choice. After creating the file, I opened it in VLC player and compared it to the DVD in Windows Media Player. The DVDs used in this testing process all have a very simple structure - a main menu

[6] "Infrarecorder." GitHub. Web.

with one chapter. Clicking on the chapter allows the video to play, and these actions are retained in the ISO files, so the experience of viewing the DVD and the ISO file is essentially the same. The only issue I noticed was that the image was stretched in VLC, but adjusting the aspect ratio to 4:3 solved the problem. Uploading these files to Google Drive and NYU Box was admittedly a slow process, and it seems best to upload in batches, though one must keep an eye out for upload failures. An ISO file stored in the cloud can be downloaded at any time and viewed in VLC or a similar player.

The library may also want to maintain access copies of these files, so I created .mp4 files that can be stored on Google Drive and NYU Box and streamed there directly. I chose Handbrake, another free and open-source application licensed under the GNU General Public License Version 2. The process of creating this file from a DVD was very similar to the ISO creation process. Once the file was created, I did a quality control check and then uploaded it to cloud storage. Handbrake is capable of creating files with MPEG-2 and MPEG-4 codecs, though it is advised that access files be created only using the MPEG-4 codec because the MPEG-2 files are not compatible with as many viewing platforms, including QuickTime on OS X 10.6.8 and Windows Media Player on Windows 7.[7]

While it is relatively simple to create master and access files of the DVDs in this collection, the interactive DVDs and CDs bring their own sets of challenges. For example, an item titled *El Mundo de los Isoseño-Guaranies* (NYU2188_01) is a complex video CD. It contains HTML documents (htm), Jscript files, Adobe Acrobat documents, .jpegs, .mp3s, .movs, .dxrs, and .mpgs. There are 22 video files and 50 audio files. Other interactive CDs and DVDs in this collection have similar structures. Attempts to create disc images using several different applications all failed, and it is unclear how the data on these discs could be cloned in their

[7] Morel M. Preserving Write-Once DVDs: Producing Disk images, Extracting Content, and Addressing Flaws and Errors. *George Blood Audio Video Film*. 2014.

entirety, though I would have liked to make more attempts on a different computer, perhaps using the command-line.

When Angela acquired this collection, she was unaware that CDs like this one would be included, and she has said that they may not be worth keeping in the collection. However, these discs do contain audio and video files that might be of interest, and those could be extracted and stored separately. I did not experiment with this idea, and decisions about keeping or deaccessioning these interactive CDs and DVDs will need to be made first, but the preservation or ACM departments can give advice on the subject.

Recommendations

Short Term

- Continue efforts to secure permissions for digitization and access
- Insert all discs into a disc drive to make sure they play normally, and that there aren't any other blank discs in the collection. Contact APCOB for replacement discs if possible.
- Decide if any items should be deaccessioned. First, it is necessary to gain a better understanding of what these interactive CDs and DVDs contain, and what their original purpose was. It may be useful to discuss this with APCOB.

Mid Term

- Create master disc image files of each disc in the collection. If working at a different workstation, pay attention to the computer's operating system, as that will determine which applications or command-line tools can be utilized.
- Create database records.
- If interactive CDs and DVDs are kept in the collection, consult with the preservation department on how to proceed.

Long Term

- Work with the Archival Collections Management department to determine best digital asset management procedures and establish a workflow.
- Copy digital files to the designated storage locations.
- Run regular checksums to ensure file integrity.

Resources

"Apoyo Para El Campesino-Indigena Del Oriente Boliviano " APCOB.

"CD / CD-R and DVD-R RW Longevity Research - Research Projects - Preservation Science." Library of Congress.

"Infrarecorder." GitHub. Web.

Iraci, Joe. 2005. "The Relative Stabilities of Optical Disc Formats." *Restaurateur* 26 (2): 134-150.

Morel, Morgan. 2014. *Preserving Write-Once DVDs: Producing Disk Images, Extracting Content, and Addressing Flaws and Errors*: Library of Congress.