

CREADORES DE MEMORIA: Los archivos sonoros y audiovisuales en México

Coordinadora
Perla Olivia Rodríguez Reséndiz



CD973.2
C74M4

Creadores de memoria : los archivos sonoros y audiovisuales en México /

Coordinadora Perla Olivia Rodríguez Reséndiz. – México : UNAM. Instituto de Investigaciones Bibliotecológicas y de la Información, 2021.

x, 174 p. – (Sistemas bibliotecarios de información y sociedad)
ISBN: 978-607-30-5581-9

1. Archivos audiovisuales. 2. Archivos sonoros. 3. Patrimonio cultural – Protección. 4. Preservación digital. 5. México. I. Rodríguez Reséndiz, Perla Olivia, coordinadora. II. ser.

Diseño de portada: Sonia Wendy Chávez Nolasco

Primera edición, 2021

D.R. © UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Ciudad Universitaria, 04510, Ciudad de México

Impreso y hecho en México

ISBN: 978-607-30-5581-9

Publicación dictaminada.

Contenido

INTRODUCCIÓN	vii
Perla Olivia Rodríguez Reséndiz	
APROXIMACIÓN HISTÓRICA AL ORIGEN DE LOS ARCHIVOS SONOROS Y AUDIOVISUALES EN MÉXICO	1
Perla Olivia Rodríguez Reséndiz	
DOCUMENTOS SONOROS Y AUDIOVISUALES EN MEMORIA DEL MUNDO	17
Catherine Bloch	
ANTECEDENTES Y SITUACIÓN ACTUAL DE LOS ARCHIVOS SONOROS QUE RESGUARDAN COLECCIONES ANTROPOLÓGICAS Y ETNOMUSICOLÓGICAS EN MÉXICO	33
Benjamín Muratalla	
CRÓNICA DE UN RESCATE: LA MÚSICA COMPUESTA PARA CINE EN MÉXICO (1956-1979)	47
Sibylle Hayem	
EL ACERVO SONORO DE RADIO UNAM. DESAFÍOS PARA SU PRESERVACIÓN SUSTENTABLE	65
María del Carmen LIMÓN CELORIO Yolanda Medina Delgado	
UN TEJIDO DE MEMORIA AUDIOVISUAL. PREGUNTAS HACIA UN ARCHIVO PARTICIPATIVO	73
María Álvarez Malvido Daniela Parra Hinojosa	
LOGROS Y RETOS DEL REPOSITORIO DIGITAL EN AUDIO DEL PROYECTO POÉTICA SONORA MX A CUATRO AÑOS DE SU CREACIÓN	85
Aurelio Meza Valdez Susana González Aktories	
EL RETO DE LA CONSERVACIÓN DE LOS DOCUMENTOS SONOROS EN MÉXICO	101
Mariela Salazar Hernández	

SITUACIÓN ACTUAL Y FUTURO DE LA PRESERVACIÓN DEL ACERVO DIGITAL EN LA FILMOTECA DE LA UNAM	123
Gerardo León Lastra	
EL ACCESO A LAS MEMORIAS FÍLMICAS. UN PANORAMA DE ACCIONES PARA EL FUTURO	143
Nila Guiss	
EL PATRIMONIO AUDIOVISUAL Y LA TELEVISIÓN EN MÉXICO: ACCESO DE CONTENIDOS DEL ARCHIVO DE TELEVISIÓN DE LA UASLP	161
Ubaldo Candia Reyna	

Situación actual y futuro de la preservación del acervo digital en la Filmoteca de la UNAM

GERARDO LEÓN LASTRA

INTRODUCCIÓN

El capítulo inicia con la descripción del acervo de la Filmoteca de la UNAM. Establece la definición de los términos *soporte digital* y *artefacto reproductor digital*. Luego se describen como métodos de trabajo de la Filmoteca de la UNAM el Control Logístico del Acervo Fílmico (CLAF) a través del cual se gestionan los activos del Acervo de la Filmoteca y del Laboratorio Cinematográfico de Restauración Digital (LCRD). Se expone con detalle el manejo físico de los activos y la verificación del estado de los soportes digitales. Como un recurso para la preservación digital, también se presentan resultados del estudio que se realizó para almacenar en la nube un espejo o copia de los soportes digitales. Finalmente, se presentan conclusiones en torno a la preservación digital sustentable de largo plazo.

EL CONTEXTO EN LA FILMOTECA DE LA UNAM

La Filmoteca de la UNAM preserva más de 400 mil rollos de película. Una parte de este acervo es patrimonio de la UNAM y otra

Situación actual y futuro...

pertenece a usuarios del servicio de depósito que ofrece la Filmoteca. El servicio de depósito en soporte fílmico analógico o digital es el medio para la custodia de los activos cinematográficos. Asimismo, la Filmoteca registra, observa y respeta lo concerniente a derechos patrimoniales y de autor. El acervo filmográfico se preserva en 16 bóvedas: ocho climatizadas con humedad y temperatura para películas con soporte en acetato de celulosa o poliéster, siete diseñadas con sistemas de prevención de incendio donde se resguardan películas con base en nitrato de celulosa y una más pequeña dedicada exclusivamente a la preservación de soportes digitales.

Desde 1989, la Filmoteca inició la catalogación de ejemplares bajo su resguardo en la llamada “Base Única de Datos” (BU-DA) que fue migrada en 2010 a intranet bajo una arquitectura cliente-servidor utilizando formas de HTML, JSP y JavaScript, Java en el *middleware* y un manejador de bases de datos SQL en el *back-end*. En noviembre de 2018 esta base de datos contenía más de 65,000 registros de largometrajes, documentales, series, colecciones, etc. El modelo de esta base de datos es un catálogo de ejemplares que dista mucho del modelo de referencia para catalogación de obras cinematográficas propuesto por la FIAF (Fairbairn *et al.* 2016), y su metodología poco flexible para reubicar ejemplares fílmicos en las bóvedas, complicaba los diseños de un modelo híbrido para mejorar los procesos de preservación fílmica y gestionar simultáneamente ejemplares fílmicos y digitales (León *et al.* 2017). Además, dificultaba la integración de una librería de cintas LTO para automatizar algunos procesos del sistema CLAF que más adelante se mencionan.

MARCO REFERENCIAL: SOPORTE DIGITAL Y ARTEFACTO REPRODUCTOR DIGITAL

En esta sección se establecen dos términos que después nos serán de utilidad: documento digital y artefacto reproductor digital (DPC 2018).

El documento digital

Todo tipo de activo o documento digital –textos, imágenes, sonidos, páginas web, películas, etcétera, incluso programas de cómputo, bases de datos, sistemas operativos y más– se codifica en colecciones de bits¹ organizadas bajo alguna convención o norma. Su interpretación requiere de una máquina virtual (programa de cómputo ejecutado en una computadora de propósito general) o de un artefacto digital específicamente diseñado para tal propósito. Estas colecciones de bits, además de que pueden transmitirse por canales de comunicación digital como Internet, o fluir por los componentes y circuitos de una computadora para su despliegue o proyección, pueden copiarse íntegramente y almacenarse en diferentes medios físicos como discos ópticos, memorias USB o “discos” SSD en estado sólido, discos o cintas magnéticas, etc., que genéricamente podemos llamar memorias digitales que se organizan mediante algún tipo de sistema de archivos y carpetas, que igualmente es codificado en bits.

De lo anterior, se desprende como corolario que:

- El documento digital no es el medio o memoria en que se almacena sino su codificación en *bits*.
- Que el documento digital únicamente adquiere valor cuando se puede interpretar para reproducirse en otra codificación perceptible e interpretable por el ser humano,
- Que esta reproducción requiere de otros dispositivos como computadoras, ejecutar máquinas virtuales, circuitos digitales de propósito específico y periféricos de algún tipo de salida como papel, pantallas, proyectores de cine, amplificadores con bocinas de audio, etcétera.
- Finalmente, la reproducción de documentos “nacidos” digitales o documentos digitalizados, tiene que ser “suficientemente cercana al original” para considerarse válida.

1 El bit es la unidad básica de información, corresponde a un dígito binario y sus valores son 0 o 1.

El artefacto reproductor digital

Entonces, y en resumen, la preservación digital demanda dos elementos que requieren conservarse:

- **Soporte digital**, esto es, los bits que codifican el documento digital, organizados con algún formato en uno o varios archivos digitales,
- **Artefacto reproductor digital**, como pueden ser una máquina física con circuitos digitales de propósito específico (p. ej. un reproductor de títulos cinematográficos Bluray con un despliegue, un amplificador y bocinas, o un reproductor de libros digitales como Kindle), o una computadora de propósito general con sus periféricos y el programa o máquina virtual que permitan la correcta presentación del documento digital orientada a la percepción humana.

Un poco de polémica

Cuando por primera vez introdujimos el término de *soporte digital* en la Filmoteca, se suscitó una polémica con algunos colegas que argumentaban que el soporte es el medio físico en que se fijan los *archivos digitales*. Entonces, para terminar con la polémica, expusimos el ejemplo de un depósito en que ingresan a la Filmoteca ejemplares digitales de obras cinematográficas en disco duro portátil. Puesto que el depositante otorga la custodia de sus activos a la Filmoteca y ésta no custodia discos en depósito, al depositante se le regresa su disco después del copiado o “ingesta” de sus archivos a nuestro sistema CLAF. Al futuro y en el momento que el depositante solicite sus activos, se le entregará una copia íntegra² y, de requerirse, eliminaremos sus archivos de nuestro sistema. Además, esta última entrega bien puede hacerse en disco duro, en cinta magnética, por Internet, etcétera.

2 Decimos que la copia del soporte digital es *íntegra* cuando cada uno de los bits de la copia es idéntico a los de su fuente.

CONTROL LOGÍSTICO DEL ACERVO FÍLMICO (CLAF)

En esta sección, primero citamos algunos antecedentes que nos condujeron al desarrollo del sistema CLAF, que además de ser un sistema para la gestión de activos digitales (MAM por sus siglas del inglés), resultó *ad-hoc* para también gestionar el acervo fílmico. Después, abordamos con detalle el manejo físico de los soportes fílmicos y digitales, y el manejo lógico, es decir, nuestra metodología para la preservación de los soportes digitales. Otros aspectos de nuestro sistema, como aquellos que se refieren al más reciente modelo de referencia de la FIAF (Fairbairn *et al.* 2016) y a la preservación fílmica, podrán consultarse en (León *et al.* 2017).

Circunstancia y motivaciones del desarrollo de CLAF

Cuando se emprendió el proyecto de un laboratorio para digitalizar, restaurar, difundir y preservar digitalmente los contenidos en película cinematográfica, obtuvimos suficientes recursos financieros para integrar una infraestructura productiva para los objetivos del proyecto al corto y mediano plazos. Este laboratorio, llamado Laboratorio Cinematográfico de Restauración Digital (LCRD), desde su integración inicial, concluida el último trimestre de 2014, cuenta con:

- Dos escáneres de película cinematográfica.
- Tres estaciones de trabajo con GPUS de alto rendimiento.
- Dos unidades de lectura/escritura de cintas LTO-6 enlazadas a un servidor dedicado a los trabajos de lectura/escritura y cálculos iniciales del monitoreo de preservación digital por *fixity check sum*.
- Un *cluster* de dos servidores ejecutando un hipervisor de virtualización de sistemas operativos, en donde se ejecutan los sistemas de información para la gestión del acervo de la Filmoteca, páginas web, desarrollo de sistemas y otras aplicaciones.
- Un servidor con arquitectura SAN para almacenamiento masivo en 120 TB crudos.

Situación actual y futuro...

- Una red local de 10 Gbps, donde está enlazado todo el equipo antes enlistado.
- Varias herramientas de *software* para restauración de imagen, sonido y color, edición y transcodificación de formatos de video digital, lo que incluye la generación de DCP para proyección en salas de cine digital.

Los fondos otorgados para la integración de este laboratorio fueron insuficientes para adquirir la totalidad de componentes y dispositivos planteados en su concepción original en vistas a su operabilidad a un plazo más largo. Aquella circunstancia obligó a un análisis de alternativas para encontrar una solución intermedia que permitiera alcanzar el objetivo de un laboratorio cumpliendo con los requerimientos inmediatos de la Filmoteca.

Un conjunto de componentes de los que se decidió prescindir estaba dirigido a disminuir los tiempos de procesos del laboratorio, otro componente, tampoco adquirido, soportaría la automatización del monitoreo del estado del acervo digital para preservarlo y, otro más, el sistema MAM, que apoyaría la gestión del acervo digital que produciría el laboratorio.

Los componentes de los que se prescindió para disminuir tiempos de procesos fueron:

- Una red local al menos 4 veces más rápida que la finalmente instalada.
- *Hardware* acelerador para corrección de color en 16 bits por componente cromática.
- *Hardware* acelerador para transformación entre múltiples de formatos digitales.
- Un proyector de cine digital para pruebas de control de calidad de DCP.

La ausencia de estos componentes obligó a trabajar casi la totalidad de los proyectos de digitalización en resolución 2K y a sólo utilizar 4K en contados proyectos de restauración digital que lo ameriten. Para hacer pruebas de proyección de DCPs, se recurre al apoyo de las salas de cine digital del CCU y de la ENAC.

El monitoreo del estado de un acervo digital para su preservación consiste en detectar aquellos ejemplares que sufren algún tipo de deterioro y debe realizarse de manera periódica. Este monitoreo se conoce como *fixity check sum*, y más adelante detallaremos como se realiza. Por ahora, es suficiente comentar que su automatización era prescindible al iniciar la creación del acervo digital de la Filmoteca, pero se ha tornado crítico conforme el acervo ha crecido y alcanzado grandes volúmenes de datos.

Por último, decidimos que, en lugar de adquirir un sistema MAM, que es imprescindible para la gestión y mantenimiento de un acervo digital, el sistema se desarrollaría en la Filmoteca. De consideración fueron las implicaciones que tal adquisición trae consigo, como su costo directo, el costo de recursos humanos para su puesta a punto acorde a su entorno operativo final y sus costos de mantenimiento. Para apuntalar la decisión, se analizó y consideró viable extender los alcances de un desarrollo iniciado un año antes. Este desarrollo llamado CLAF, originalmente estaba dirigido a sustituir el obsoleto sistema BU-DA para responder a antiguos requerimientos de sus usuarios y de auditoría interna. Finalmente, CLAF resultó en un MAM *ad-hoc* que permite gestionar activos filmicos y digitales de manera unificada, cuyos soportes físicos se identifican con códigos de barras. El desarrollo se planteó en fases, que fueron entrando en operación a partir de abril de 2015 en el siguiente orden cronológico: registro de nuevos ingresos filmicos, gestión de proyectos de digitalización y del acervo digital, gestión de préstamos filmicos y digitales, levantamiento periódico del inventario físico del acervo y, por último, un módulo para migrar paulatinamente los registros de ejemplares filmicos en el antiguo sistema BU-DA hacia el nuevo sistema CLAF. Puesto que esta migración requiere la adhesión de códigos de barras a las latas de filme, se ideó la estrategia de migrar gradualmente de manera simultánea con las tareas de limpieza y mantenimiento del estado físico de las películas, tareas que invariablemente se realizan antes de proyectar o digitalizar una película y que son indispensables para la preservación de sus contenidos.

Los procesos del laboratorio cinematográfico de restauración digital (LCRD)

En el último trimestre de 2014 la Filmoteca comenzó a operar³ su Laboratorio Cinematográfico de Restauración Digital (LCRD), en donde, utilizando dos escáneres se digitalizan el sonido y la imagen registrados en película cinematográfica de 8, 16 y 35 mm positiva o negativa. El LCRD cuenta con infraestructura de *hardware* y *software* para restaurar imagen y sonido digitales, corregir color, transcodificar a múltiples formatos digitales, difundirlos en salas de cine con proyección digital o prepararlos para su difusión por Internet o en discos ópticos. Los formatos digitales se preservan en cintas magnéticas Ultrium LTO de sexta generación, cuyos metadatos se registran en CLAF. En la intranet de Filmoteca, los contenidos y metadatos se gestionan con CLAF que también soporta la reproducción en línea –del *proxy* en baja resolución– de los ejemplares digitalizados.

Las películas típicamente se digitalizan en resolución 2K con 10 bits de profundidad por componente cromática, lo que se traduce en un consumo por encima de los 10 MB por fotograma no comprimido; el sonido se digitaliza en PCM, envuelto en WAV, con muestras de 24 bits. Para trabajos de restauración digital, el muestreo del sonido puede incrementarse a 96 kHz y la imagen digitalizarse en resolución 4K con 16 bits de profundidad por componente, lo que resulta en cerca de 74 Mb por fotograma y consume hasta 9 TB para un largometraje de 90 minutos.

Organización e identificación física del acervo

La estantería en todas las bóvedas de la Filmoteca está identificada con un código de barras en cada *ubicación* (ver imagen 1). Una

3 Desde el inicio se calculó el *fixity check sum* de los ejemplares digitales resguardados en LTO. En tanto que CLAF no era operable, los metadatos de los primeros ejemplares digitalizados se capturaron unos meses en tres hojas relacionales de Excel, sus datos después se migraron a CLAF.

ubicación es un volumen claramente delimitado en la estantería por postes, charolas o cualquier tipo de separador; para efectos prácticos, cada ubicación tiene una capacidad para almacenar entre 10 y 40 latas con rollos de película o estuches con cintas.

Imagen 1. Películas del acervo en latas.



Imagen 2. Acervo digital en LTO.



Para localizar fácilmente una ubicación, su identificador en código de barras (ver imagen 2) está organizado a modo de coordenadas espaciales con la siguiente sintaxis:

Situación actual y futuro...

bóveda(2)+nivel(1)+número de anaquel(2)+cara del anaquel(1)+número de columna del anaquel(2)+un guion separador(1)+número de entrepaño(2)+un espacio separador(1)+un dígito verificador(1)

Entre paréntesis aparece el número de caracteres que ocupa la entidad cuyo nombre le precede. El signo de “+” simboliza “seguido de” o concatenación. Utilizamos dos caracteres para la bóveda. El nivel es necesario porque hay bóvedas con más de un piso. El código permite tener hasta 100 anaqueles por piso, y todo anaquel tiene una o dos caras que se identifican con la letra “A” o “B”. Las columnas y entrepaños de los anaqueles se numeran con dos dígitos.

Por otro lado, cada contenedor -es decir cinta LTO y su estuche en el caso digital, y lata en el caso fílmico- se identifica con un secuencial irrepetible dividido en dos secuencias. La secuencia para cintas y estuches de LTO inicia en 1 y la secuencia para filme inicia en 1048575. Previendo que algún día nos haremos de una librería de cintas, el identificador de LTO se arma con su secuencial y otros caracteres de acuerdo con la especificación en (IBM 2009) del consorcio Ultrium LTO (Ultrium LTO, 1999-2018).

Cuando un contenedor se coloca en alguna bóveda, el operador de CLAF relaciona el identificador del contenedor con el identificador de su ubicación. Para las tareas de ubicación y de recolección de materiales, se utiliza un dispositivo móvil con lector de código de barras integrado (ver imagen 3).

Imagen 3. Ubicación y recolección de contenedores.



Preservación lógica del soporte digital o *fixity check sum*

Jefferson Bailey define: “[...] en el contexto de la preservación, *fixity* significa tener la certeza de que un archivo digital permanecerá inalterado, es decir, fijo” (Owens 2014, s.p). El término *fixity check sum* se utiliza para referirse a técnicas de cifrado y criptografía aplicadas a la preservación del soporte digital (CCSDS 2012). En esta sección se expondrá a detalle cómo trabajamos el *fixity check sum* del acervo digital.

Cada proyecto de digitalización de una película puede tomar días y se gestiona utilizando CLAF. Metódicamente, todo proyecto resulta en una nueva arborescencia de archivos y carpetas alojada en el almacenamiento masivo. Su raíz se identifica con un prefijo numérico único y un sufijo que corresponde al título de la obra cinematográfica. Cada subcarpeta en el segundo nivel contiene formatos digitales diferentes del mismo ejemplar de la película; estas carpetas se nombran con acrónimos basados en el tipo de formato o envoltura de formato que contienen (DPX, WAV, DCP, MOV, PROXY, etc.). La versión *proxy* en baja resolución invariablemente se genera y se copia a una partición del almacenamiento en donde queda disponible para su reproducción en línea por los usuarios de CLAF. Los metadatos de catalogación de cada ejemplar digital y de su formato se registran en CLAF.

Concluido el proyecto, iniciamos el proceso de preservación o *fixity check sum*, utilizando dos programas o *scripts* que desarrollamos con CMD de MS-Windows. El primer *script* se utiliza para copiar de disco a cinta LTO y verificar la integridad de la copia resultante, y el segundo, se utilizará periódicamente para monitorear que el contenido de cada cinta no ha sido alterado. Para el cálculo del *fixity check sum* y su registro en .XML, ambos *scripts* se apoyan en el paquete FCIV disponible de forma gratuita (Microsoft 2018). En seguida describimos estos dos *scripts*.

Tomando como entrada el nombre de la raíz de un proyecto alojado en el almacenamiento masivo, el *script* de copiado y verificación del copiado, hace lo siguiente:

Situación actual y futuro...

1. Calcula y almacena en disco un archivo .XML que contiene el nombre y el *fixity check sum* de cada uno de los archivos del proyecto empleando el algoritmo MD5 (Wikipedia 2018a).
2. Copia todos los archivos del proyecto a cinta LTO.
3. Leyendo los archivos ya copiados en la cinta LTO, se repite el cálculo de MD5 registrándolos en otro archivo .XML también en disco.
4. Finalmente, se comparan los dos archivos XML y, si el *script* no detecta diferencias, concluye copiando el XML a la cinta a modo de *metadatos de preservación* de los ejemplares del proyecto contenidos en la cinta.

De manera periódica debe ejecutarse el segundo *script* para monitoreo y preservación, que toma como entrada una cinta LTO del acervo y hace lo siguiente:

1. Copia a disco los *metadatos de preservación* alojados en el XML en la cinta.
2. Al leer todos los demás archivos en la cinta vuelve a calcular sus códigos MD5 registrándolos en un nuevo XML.
3. Compara los dos archivos XML y, si no se detectan diferencias, concluye que los archivos en cinta siguen sin alteración en cada uno de sus bits.
4. Se concluye actualizando la fecha de la última revisión de la cinta en CLAF.

En la explicación anterior, por claridad, omitimos que: a) en una misma cinta generalmente copiamos dos o más proyectos, b) si el proyecto supera la capacidad de una LTO, lo partimos en dos o más raíces diferentes utilizando más de una LTO, con su correspondiente registro relacional en CLAF. Cabe comentar que generamos dos copias idénticas de cada cinta, una ubicada en la bóveda digital y la otra programada para ubicarse en otra bóveda remota. Ya que se tienen las dos copias en LTO de un proyecto, éste se elimina del almacenamiento masivo.

Con la actual infraestructura del LCRD los procesos de cálculo, copiado y verificación del *fixity check sum*, toman cerca de 30 horas por cinta. Aunado a las cargas de trabajo cotidianas del LCRD tenemos que reconocer que el monitoreo de preservación no lo estamos aplicando periódicamente, pues es imposible con los recursos humanos y de *hardware* disponibles. La noticia esperanzadora es que el monitoreo es automatizable mediante la integración a nuestro sistema de una librería de cintas LTO robotizada, lo que, conforme crece el acervo digital y envejecen las citas que lo contienen, se torna cada vez más urgente su integración. Además, la eventual migración del acervo en LTO-6 a nuevas tecnologías de almacenamiento, apoyada en una librería de cintas robotizada, es también automatizable, y es indispensable para eliminar el posible error humano y las consecuentes pérdidas de ejemplares digitales.

En abril de 2020, una sola copia del acervo digital de la Filmoteca ya acumulaba 411 Tb en 164 cintas LTO-6, lo que representa poco más de 35 millones de fotogramas provenientes de 1248 títulos digitalizados. Antes que la generación 6 de LTO se torne insostenible, eventualmente el acervo digital de la Filmoteca tendrá que migrarse a otro soporte físico del soporte digital, que podrá ser una nueva generación de LTO o almacenamiento SSD.

LA NUBE Y EL ARTEFACTO REPRODUCTOR DIGITAL

En esta sección, se presentan los resultados de un estudio de costos para mantener un espejo del acervo en la nube y se cita una iniciativa esperanzadora dirigida, entre otros objetivos, a la preservación del artefacto reproductor y por ende a la posible preservación al muy largo plazo de los acervos de documentos digitales en general y el de la Filmoteca en particular.

El espejo en la nube

A invitación de la Coordinación de Asesores del Rector de la UNAM, y considerando que entre más copias de un acervo se tengan más

Situación actual y futuro...

disminuye la probabilidad de su daño o pérdida, en 2018 hicimos una proyección de costos para almacenar en la nube un espejo del acervo digital de la Filmoteca para el período 2019-2024 (CoNTI 2018).

Estimamos que para 2024, el tamaño de una sola copia del acervo crecerá a 940 TB, casi un petabyte. El costo total *estimado* a seis años, por servicios de almacenamiento en la nube, resultó un 33 por ciento menor que los costos por el mantenimiento de infraestructura del LCRD (que incluyó para este estudio dos escáneres de película cinematográfica, almacenamiento masivo SAN de 80 TB en raid-5, un clúster de dos servidores con virtualización, un servidor con dos *drives* LTO-6, todos comunicados por un *switch* de red local de 10 Gbps, un UPS para mantenerlos y una manejadora de aire y humedad acondicionados). Esta estimación asumió un incremento anual compuesto de 15 por ciento por costos de mantenimiento de infraestructura por empresas locales, y de 5 por ciento por costos por servicios en el extranjero. Luego, almacenar el espejo en la nube resulta muy atractivo para un escenario en que es posible bajar significativamente los costos *on-premises* (Wikipedia 2018b).

Desafortunadamente, para el caso de la Filmoteca, el ahorro de costos en mantenimiento *on-premises* bien identificados para este espejo es de apenas un 5 por ciento, pues no existen servicios de digitalización de filme en la nube. Además, en los porcentajes anteriores no se consideró que mover en 2024 la totalidad del espejo en la nube hacia alguna otra instalación, representaría un 72 por ciento del costo total acumulado por servicios en la nube. Entonces, por el momento, consideramos que no es costeable para la Filmoteca mantener un espejo en la nube.

Las estimaciones anteriores no consideraron costos por consumo de energía eléctrica, ni costos de procesamiento del *fixity check sum* en la nube ni *on-premises*, que es imprescindible, ni resolvieron ni costearon la solución del establecimiento de las referencias en la base de datos y metadatos de CLAF hacia los activos digitales en la nube.

Preservación del artefacto reproductor digital

Como antes se mencionó, hay dos grandes clases de artefacto reproductor, una constituida por los de *hardware* puro (como es un reproductor de Bluray, con un monitor sencillo de televisión y amplificador de sonido) y la otra por artefactos compuestos por *hardware* y *software* (como son los celulares “inteligentes”, las computadoras de escritorio y los proyectores de DCP) que invariablemente incluyen una computadora de propósito general. En el caso de un acervo digital como el de la Fílmoteca, utilizamos artefactos pertenecientes a la segunda clase.

Aunque es de esperarse que la evolución tecnológica del audiovisual digital imponga aumentos en resolución espacial, espectral y temporal, que tornarán obsoletos los formatos del audiovisual digital actual, nuestros órganos perceptuales no evolucionarán tan vertiginosamente. Entonces y en teoría, podremos reproducir nuestro acervo digital de sonido e imagen en movimiento con la mera preservación del *software* y su necesaria adaptación al *hardware* por venir. Como soporte a esta afirmación, cito ahora varios conceptos expresados en (Abramatic *et al.* 2018):

- El *software* es un mediador *esencial* para acceder a *cualquier* tipo de información digital.
- Una creciente parte de nuestro *conocimiento* está constituida por (o depende de) artefactos de *software*.
- Hoy es bien aceptado en la extensa literatura sobre la preservación digital que la preservación al largo plazo requiere acceso irrestricto al código fuente de las herramientas utilizadas para lograr su objetivo.
- Que lo anterior incluye tanto al *software* libre y abierto, como al propietario y cerrado.
- Sin embargo, no hemos cuidado con la entereza requerida esta valiosa forma del conocimiento que es el *software*.
- Los científicos y los profesionales de cómputo tienen el deber, la responsabilidad y el privilegio de armar la infraestructura de un archivo universal de código fuente.

Situación actual y futuro...

Abramatic y otros autores promueven la organización “Herencia del Software” dedicada a preservar todo el *software* (Abramatic *et al.* 2018). Esta organización nació como iniciativa del Instituto Nacional de Investigación en Informática y Automatización de Francia, y tiene acuerdos con la Unesco y grandes consorcios de la industria (Software Heritage 2018).

Ante el escenario universal actual se nos ocurre que organizaciones como la Filmoteca de la UNAM deberían recurrir estrictamente a herramientas de *software* abierto. En nuestro caso así lo hemos hecho en lo que se refiere a los sistemas de información que recurren a la plataforma Linux y otras herramientas de desarrollo de *software* y bases de datos, aunque sobre máquinas virtuales apoyadas en una plataforma de virtualización propietaria. Tampoco ha sido posible recurrir a *software* abierto de herramientas de restauración digital, edición de video digital, transcodificación y DCPS, el principal obstáculo ha sido la carencia de recursos humanos para esta tarea.

Por otro lado, es muy probable que los servicios en la nube evolucionarán para proveer herramientas de *software*, propietario o libre, para proyectos de restauración y preservación digital. Cito un ejemplo apoyando esta especulación: cuando realizábamos las pruebas de proyección de costos para mantener el acervo en un espejo alojado en la nube, mi colaborador, el Ing. Gustavo Lucio José, probó el cambio de formato o transcodificación del DCP 4K de un largometraje a formato ProRes 4:2:2 HQ Full HD; para nuestra sorpresa, nos percatamos que esta prueba resultó muy económica, eficiente y de calidad.

CONCLUSIONES

Después de explicar cómo integramos un laboratorio de digitalización con el desarrollo de CLAF como herramienta para la gestión y preservación de los productos del laboratorio, mostramos una proyección del crecimiento del Acervo de la Filmoteca para 2024. Lo que deja prever el correspondiente aumento de las presiones financieras para la preservación de este acervo en soporte digital.

Ante nuestro compromiso con la Filmoteca de buscar alternativas para preservar al muy largo plazo su acervo, presentamos los resultados de un estudio de la Filmoteca que compara los costos de un espejo del acervo en la nube con los costos de mantenimiento *on-premises*, en una proyección de 2018 a 2024. Y, aunque por lo pronto no resulta costeable el espejo en la nube, cuando llegue el momento de una eventual migración del acervo digital en su estado actual y/o de una renovación del equipamiento del LCRD, los servicios en la nube deberán ser evaluados como alternativa, aún más si para entonces habrán aumentado las velocidades de transferencia y disminuido sus costos por la proliferación de la competencia de mercados.

Finalmente, se planteó como esperanzadora la iniciativa del INRIA para la creación del repositorio universal “Herencia del Software”, como medio sustentable para la preservación del artefacto reproductor de documentos en soporte digital.

AGRADECIMIENTOS

Este escrito no hubiese sido posible sin la confianza depositada en el grupo de trabajo del autor, por parte del licenciado Hugo Villa Smythe, Director de la Filmoteca de la UNAM, y por su antecesora la licenciada María Guadalupe Ferrer Andrade. El autor hace patente su agradecimiento a los ingenieros Manuel Comi Xolot, Luis Felipe Maciel Mercado y Gustavo Lucio José por su disciplina, aportaciones y dedicación al trabajo colaborativo que ha fructificado en los resultados que este capítulo reporta.

BIBLIOGRAFÍA

- Abramatic, Jean-François, Roberto Di Cosmo, and Stefano Zacchiroli. 2018, “Building the Universal Archive of Source Code: A global collaborative project for the benefit of all”. *Communications of the ACM*, (Vol. 61, No. 10, October 2018).

Situación actual y futuro...

- CCSDS. Consultative Committee for Space Data Systems. 2012. "Reference Model for an Open Archival Information System (OAIS), Recommended Practice, CCSDS 650.0-M-2, Magenta Book". Estados Unidos: Management Council of the Consultative Committee for Space Data Systems. <https://public.ccsds.org/pubs/650x0m2.pdf>.
- CoNTI. Coordinación de Nuevas Tecnologías e Informática. 2018. "Reporte sobre costos para mantener una copia espejo en la nube del acervo digital de la Filmoteca de la UNAM. Reporte interno de la Dirección General de Actividades Cinematográficas, UNAM, noviembre de 2018.
- DPC. Digital Preservation Coalition. 2018. "Digital Preservation Handbook". <https://www.dpconline.org/handbook>.
- Fairbairn, Natasha, Maria Assunta Pimpinelli, Thelma Ross. 2016. "FIAF Moving Image Cataloguing Manual", edited by Linda Tadic, project by Nancy Goldman for the FIAF Cataloguing and Documentation Commission. <http://www.fiafnet.org/images/tinyUpload/E-Resources/Commission-And-PIPResources/CDC-resources/20160920%20Fiaf%20Manual-WEB.pdf>.
- IBM Corporation. 2009. "IBM LTO Ultrium Cartridge Label Specification", Part Number 19P0034 EC - M10321 Feb. 2009. <http://www-01.ibm.com/support/docview.wss?uid=ssg1S7000429&aid=1>.
- León Lastra, Gerardo, Manuel Comi, Gustavo Lucio, Luis F. Maciel. 2017. "Sistematización del Acervo Fílmico de la Filmoteca de la UNAM". En Rodríguez, Perla y Munera Fernández. *Conectando los saberes de bibliotecas, archivos y museos (BAM) en torno a la preservación de documentos analógicos y de origen digital*. México: IIBI-UNAM 2019.
- Microsoft. 2018. "Availability and description of the File Checksum Integrity Verifier utility". <https://support.microsoft.com/en-us/help/841290/availability-and-description-of-the-filechecksum-integrity-verifier-u>.
- Owens, Trevor. 2014. "Protect Your Data: File Fixity and Data Integrity", Jefferson Bailey. <https://blogs.loc.gov/thesignal/2014/04/protect-your-data-file-fixity-and-data-integrity/>.

- Software Heritage Organization. 2018. <https://www.softwareheritage.org/?lang=es>.
- Ultrium LTO. Consorcio Ultrium LTO. 1999-2018. <https://www.lto.org/>.
- Wikipedia. 2018a. “MD5”. <https://es.wikipedia.org/wiki/MD5>.
- . 2018b. “On-premises software”. https://en.wikipedia.org/wiki/On-premises_software

Creadores de memoria: Los archivos sonoros y audiovisuales en México. Instituto de Investigaciones Bibliotecológicas y de la Información/UNAM. La edición consta de 100 ejemplares. Coordinación editorial, Anabel Olivares Chávez; revisión especializada, Carlos Ceballos Sosa; corrección de pruebas, Carlos Ceballos Sosa; revisión de pruebas, Valeria Guzmán González; formación editorial, Sonia Wendy Chávez Nolasco. Fue impreso en papel cultural de 90 gr en los talleres de Dataprint, Georgia 181, Col Nápoles, Alcaldía Benito Juárez, Ciudad de México, C.P. 03810. Se terminó de imprimir en octubre de 2021.