

Factores tecnológicos, legales y documentales de la preservación documental digital

Juan Voutssas M. *

*Artículo recibido:
8 de enero de 2009.*

*Artículo aceptado:
21 de octubre de 2009.*

RESUMEN

Se analiza la problemática actual de la enorme producción y acumulación mundial de información en forma de documentos electrónicos o digitales, así como los problemas derivados del riesgo y pérdida de esa información y las diversas estrategias para su posible preservación. Se estudian y establecen con detalle los factores que inciden a favor y en contra de la preservación de documentos digitales a largo plazo; en particular los factores tecnológicos, legales y documentales.

Palabras clave: Bibliotecología; Ciencias de la Información; Bibliotecas y archivos digitales o electrónicos; Preservación documental digital;

* Centro Universitario de Investigaciones Bibliotecológicas de la UNAM, México. voutssas@servidor.unam.mx

Factores de la preservación digital; Tecnológicos;
Legales; Documentales.

ABSTRACT

Technological, legal and documentary factors about digital documental preservation

Juan Voutsas M.

Today's problematic about the huge world production and accumulation of information conformed as digital or electronic documents is analyzed, as well as the problems related to the risk or possible loss of such information and the different strategies for insuring its preservation. Factors which work in favor or against long term preservation of digital documents are studied and established, particularly those of technological, legal or documentary nature.

Keywords: Library Science; Information Science; Digital or Electronic Archives or Libraries; Digital documentary preservation; Digital preservation factors; Technological; Legal; Documentaries.

I. INTRODUCCIÓN

En el año 2000, Lyman y Varian publicaron un estudio en el que trataron de establecer el tamaño de la información digital original que el mundo producía por año. En ese entonces estimaron que el total producido ese año era de dos exabytes de información original, esto es dos trillones de bytes, o 10^{18} (diez a la décimo octava potencia) [Lyman y Varian, 2000].¹ En 2003 actualizaron su estudio y afirmaron que la cantidad era ya de tres a cinco exabytes [Lyman y Varian, 2003].² Al margen del error que pudiesen tener estos estudios, nos permiten darnos una idea aproximada del la cantidad de información digital que el mundo produce cada año en nuestro tiempo, ya

1 Peter Lyman y Varian, Hal. 2000. "How Much Information?", en *Journal of Electronic Publishing*, Diciembre, 2000, vol. 6:2, ISSN 1080-2711, disponible agosto 2008 en: <http://www.press.umich.edu/jep/06-02/lyman.html>

2 Peter Lyman y Varian, Hal. 2003, "How Much Information? 2003", en Sitio Web Oficial de la Universidad de California en Berkeley, EUA, disponible agosto 2008 en: <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>.

que en general ignoramos su verdadera y enorme dimensión. Además de su formidable tamaño es importante considerar que esta información se acumula y crece en forma geométrica. La mayor parte de esa información es de tipo informal o personal –correos electrónicos, páginas, mensajes y fotografías familiares, etcétera– y no está relacionada directamente con el mundo de las bibliotecas o archivos, pero es innegable que el remanente de esa información todavía es una cantidad muy considerable, y sí pertenece al ámbito de lo que a estas organizaciones les es deseable y necesario preservar.

Por lo mismo en la actualidad la preservación documental ya no está relacionada solamente con el mundo del papel u otros documentos tangibles en soportes “tradicionales”, sino también con los documentos electrónicos. Las bibliotecas, conforme tienden a su nueva organización en forma de bibliotecas digitales, toman las medidas para preservar y distribuir tanto las colecciones que estuvieron originalmente en soportes tradicionales y ahora se convierten en digitales, como las colecciones producidas originalmente en formatos digitales. Y no tan sólo las bibliotecas; todas las instituciones que tienen como misión la preservación documental enfrentan la misma problemática: museos, archivos, sociedades históricas, etcétera.

Quienes no conocen el problema a fondo piensan que se trata simplemente de pasar todos los documentos a través de un escáner, cámara digital o cualquier otro dispositivo de digitalización, y que así se obtendrán fácil y masivamente miles y miles de documentos digitales que pueden de esta forma ser distribuidos vía la red. Si los documentos nacieron originalmente digitales, mejor: sólo se trata de almacenarlos en un computador para ser ulteriormente distribuidos. Nada más lejos de la realidad: el proceso de ofrecerles a las comunidades de usuarios colecciones documentales estructuradas y de preservarlas para futuras generaciones, implica retos formidables que los bibliotecarios, archivistas, y muchos otros profesionales de la información deben poder afrontar y resolver, y para ello deben saber cómo. Y es esencial que existan los medios para identificar y mantener *para siempre* la información registrada. Sin ella no habría registro del pasado, no habría manera de identificar, entender o reproducir la forma de vivir de la sociedad de un momento dado –en este caso nuestra sociedad contemporánea–, ni los estudios y descubrimientos recopilados en el pasado con objeto de hacer avanzar el conocimiento.

Quienes ya han recorrido este camino de la preservación digital saben por experiencia que es tortuoso, empinado y nada fácil. El proceso de digitalización documental y su preservación para uso posterior es un reto que implica amplios conocimientos tanto teóricos como metodológicos: decisiones delicadas; fuertes inversiones en tecnología y en recursos humanos calificados;

serias limitaciones legales; obsolescencia de dispositivos; reconversión de tiempo en tiempo y, por sí todo ello fuera poco, grandes riesgos. Muchos responsables de acervos han postergado su decisión de entrar a este mundo de la preservación digital en espera de que llegaría un momento en que la tecnología se volviese estable y los costos se volvieran competitivos: vana esperanza. Es cierto que la tecnología y sus costos mejoran cada día, pero en cambio, la cantidad y variedad de materiales que hay que digitalizar crece en forma exponencial, lo cual desvanece las ventajas así obtenidas. No es recomendable bajo ningún aspecto esperar a que los parámetros se estabilicen y los riesgos desaparezcan. Eso no está sucediendo. Es necesario preservar ya y no hay opciones que omitan o le den vuelta a este camino tecnológico. Por lo tanto hay que entender el fenómeno ahora, prepararnos hoy para el proceso y los riesgos con objeto de hacerlo bien desde el principio, sin costos innecesarios y sin errores que después sea necesario corregir –o peor aún–, lamentar.

Muchos bibliotecarios y archivistas profesionales tienden todavía a menospreciar el medio electrónico. Bajo la premisa de “lo que no está impreso no está publicado”, simplemente tienden a ignorar el material digital en la red por considerarlo fútil. Variados autores han establecido ya claramente que la Web, como un todo, no es una biblioteca digital en sí misma, Gorman³ y Lynch⁴ por ejemplo –y estoy totalmente de acuerdo con ellos–, pero también es un hecho incuestionable que *todas las bibliotecas digitales forman parte de la Web*, así como una muy considerable parte de los archivos digitales. Y si todas las bibliotecas y esos archivos digitales forman parte de la Web, millones de documentos perfectamente válidos forman ya parte de acervos de bibliotecas y archivos, ahora electrónicos. Pero esos millones de documentos formales –y digo formales porque forman parte del acervo de esas colecciones digitales–, así como muchos millones más de documentos “informales” se crean y se pierden con sobrecogedora frecuencia.

Cuando un documento se digitaliza para formar parte de una colección, obedecemos por lo general a uno de los dos motivos básicos por los cuales digitalizamos nuestras colecciones: preservar o distribuir el material documental. Por un lado muchas bibliotecas, archivos, etcétera tienen la misión de preservar las colecciones documentales para futuras generaciones. El propósito es que el material trascienda al tiempo y pueda estar al alcance de lectores en años o siglos futuros. Por el otro lado, muchas de estas instituciones

3 Michael Gorman, 1998, “What is the Internet”, en *The One Person Library*, vol. 15, Núm. 6 (6/98) p. 5.

4 Clifford Lynch, 1997, “Searching the Internet”. en *Scientific American*, March 1997, disponible agosto 2008 en: <http://www.hackvan.com/pub/stig/articles/trusted-systems/0397lynch.html>.

u otras semejantes, si bien no tienen la función de preservar, desean hacer llegar cierto material documental a sus comunidades usuarias, cada vez a mayor número de lectores, a mayor distancia y durante todo el tiempo. Desean que los documentos se desplacen, se divulguen, se consulten ampliamente, y que lleguen a numerosos lectores y usuarios. Esa función bibliotecaria de distribuir materiales es también causa y motor de la digitalización. Por ello puede desearse entonces convertirla a un nuevo soporte para maximizar las posibilidades de que esto suceda. Se busca un soporte ágil, que esté al alcance de muchos usuarios y, por supuesto, que pueda hacerse a costos *razonables*. Se digitaliza también para facilitar la distribución o, en el caso de los archivos, el acceso y la consulta. Son estas dos razones –preservar y distribuir– las que en mayor parte hacen moverse todo el mecanismo de la digitalización y crean las bases para el trabajo en bibliotecas digitales. Nótese que la función de preservar permanece claramente estipulada en la caracterización de biblioteca digital moderna. Y lo ha estado por siempre en los archivos, digitales o no.

A pesar de la dificultad, los riesgos y los costos de preservar, es imposible soslayar el problema. Todos los grupos y conglomerados humanos, todas las regiones y países poseen riquezas documentales que necesariamente deben preservar y distribuir. Forman en esencia su patrimonio cultural; es impensable no considerarlo así y no tratar de preservarlo. Por ello se han iniciado en diversas partes del mundo una serie de esfuerzos tendientes a darle forma y estructura a ese patrimonio documental de diversas regiones y comunidades, así como a crear una memoria documental, tanto en los soportes tradicionales como en los nuevos soportes digitales. Por eso en sociedades avanzadas en las *Tecnologías de Información y Comunicación* –TIC– se observa ya claramente la preocupación de crear una memoria documental digital para la generación actual y las futuras. Este fenómeno puede verse ya claramente en la Unión Europea, Los Estados Unidos de América, los países nórdicos, Canadá, Australia, etcétera. Todas estas regiones diseñan ya estrategias que les permitan, por una parte, recolectar todo su patrimonio cultural escrito, filmado o grabado para que pueda ser preservado y distribuido vía la Red. Entre ellos podemos poner como ejemplos los proyectos *American Memory* o *Digital Preservation* dentro del *National Digital Information Infrastructure and Preservation Program* de la unión americana, *eContentplus* y *European Digital Library* –también llamada *Europeana*– de la Comunidad Económica Europea; *ARNO* –*Academic Research in Netherlands on-line*–, de Holanda e *InterPARES*, esfuerzo multinacional de preservación de archivos digitales coordinado por Canadá, sólo por mencionar algunos.

En general estos proyectos no son triviales, implican retos formidables que deben afrontar bibliotecarios, archivistas y muchos otros profesionales

de la información; factores complejos que deben ser resueltos a la hora de desarrollar este tipo de proyectos. Una lista indicativa de estos retos y factores sería:

- La enorme diversidad del material a preservar: libros, revistas, manuscritos, periódicos, fotografías, discursos, videos, películas, grabaciones, etcétera.
- La gran cantidad de material transaccional de archivos que hay que preservar: registros, certificados, actas, oficios, correos, memorandos, reportes, contratos, patentes, etcétera.
- La vastísima cantidad de material a digitalizar retrospectivamente: millones y millones de piezas documentales que se encuentran en los acervos de bibliotecas, archivos, museos, etcétera.
- La enorme cantidad de material originalmente digital y su enorme tasa de aparición y de pérdida.
- La falta de conciencia sobre el valor histórico de los acervos documentales.
- Los factores tecnológicos que inciden en la preservación de material digital a mediano y largo plazo: soporte, formatos, *hardware* y *software* asociados para su visualización, así como la previsión de obsolescencias tecnológicas.
- Seguridad de la información.
- Restricciones y consideraciones legales con respecto a derechos de autor y de acceso, así como las relativas a los derechos de privacidad.
- Inclusión de los metadatos pertinentes para la recuperación adecuada de la información.
- Integración de los tres ejes principales: digitalización, preservación y distribución en-línea.
- Mecanismos que permitan garantizar el acceso futuro a la información por parte del público. Cierre de la *brecha digital* en este aspecto.
- Y obviamente, los costos asociados con la preservación de este material digital, aparte de los generados por la producción del mismo.

La problemática que se desprende de estos proyectos es formidable; no pueden ser enfrentados a nivel de organizaciones aisladas ya que sin duda rebasaría sus capacidades; debe ser solucionada a nivel de naciones, regiones o grandes consorcios, y contar con la participación de múltiples sectores públicos y privados. En cuanto a las personas que participen en ellos, sucede lo mismo; este tipo de proyectos deben ser tratados con enfoques trans y multidisciplinarios. Se requiere del concurso de planeadores, legisladores,

ejecutivos, gobernantes, profesionales del cómputo, de la bibliotecología y la archivonomía, editores, proveedores, etcétera, para un adecuado diseño y puesta en marcha de algún plan coherente.

Muchas cuestiones deben plantearse y resolverse. De la lista que hemos presentado se desprenden ya algunas de ellas: ¿cuánto material hay para digitalizar? ¿cuál conviene digitalizar? ¿quién debe hacer qué parte? ¿cómo repartirlo? ¿cómo evitar que se pierda lo que es originalmente digital; es decir, páginas *Web*, publicaciones y comunicados electrónicos, etcétera? ¿quién es el responsable de guardarlo? ¿cuáles son los factores que incidirán en la preservación a largo plazo? Creo que muchas de estas preguntas pueden contestarse a través de la última: los factores que inciden en la preservación. Tales factores pueden agruparse de varias formas para su estudio. En lo personal, yo los he agrupado en seis categorías: *Factores tecnológicos, legales, documentales, culturales, económicos y sociales*.

En este artículo abordaremos con detalle los factores tecnológicos, legales y documentales. Previamente, debemos establecer lo que consideramos como conservación, preservación y restauración digital.

2. LOS CONCEPTOS DE CONSERVACIÓN, PRESERVACIÓN Y RESTAURACIÓN DIGITAL

En el campo de la seguridad informática se definen cuatro factores primordiales que pueden afectar a la información: amenaza, impacto, vulnerabilidad y riesgo. Para poner en contexto el problema es necesario enunciar en primer lugar, cuáles son las posibles amenazas con las que se enfrenta la documentación digital, las que trataremos de evitar precisamente a través de la preservación digital. En términos generales, podemos agrupar las amenazas en cinco grupos principales; éstos son:

- Obsolescencia tecnológica de los componentes de *hardware* o *software* necesarios para leerlos.
- Fallas de hardware, software o red en la escritura o transmisión de los documentos, que arruinen su integridad.
- Errores humanos que pudiesen arruinar el registro o almacenamiento de la información.
- Desastres naturales o ataques deliberados a la información.
- Fallas organizacionales o económicas cuando el poseedor de la información no pueda o no desee conservarlos más.

En función de lo anterior establezcamos una definición de los términos que nos interesan, en especial para los documentos y su versión digital:

- *Preservación Digital*: acciones específicas cuyo fin ulterior y a largo plazo, sería asegurar la permanencia y acceso del contenido de documentos digitales a lo largo del tiempo y las tecnologías, independientemente de su soporte, formato o sistema. Para ello, debemos mantenerlos, esto es, protegerlos y resguardarlos anticipada y permanentemente; y en caso de deterioro o daño debemos tratar de restaurarlos.
- *Conservación Digital*: acciones tomadas a corto plazo para anticipar, prevenir, detener o retardar el deterioro del soporte de obras digitales con objeto de tenerlas permanentemente en condiciones de usabilidad, tales como la estabilización tecnológica, la reconversión a nuevos soportes, sistemas y formatos digitales con el fin de garantizar la trascendencia de los contenidos.
- *Restauración Digital*: acciones para recuperar, reparar, renovar o volver a poner un documento digital en el estado, estimación o accesibilidad que antes tenía.

Para fines de este trabajo aclaramos que el objeto de estudio de este trabajo es el de analizar la preservación digital como fin ulterior y a largo plazo; para ello, debemos aprender acerca de la conservación digital; esto es, cómo crear, proteger y resguardar anticipadamente los documentos digitales para lograr la permanencia a largo plazo de sus contenidos. En caso de deterioro o daño los debemos restaurar.

3. FACTORES TECNOLÓGICOS DE LA PRESERVACIÓN

3.1 *Materiales a preservar*

A continuación debemos establecer el alcance del objeto de estudio; esto es, la tipología de materiales que serán susceptibles de preservarse. Ello depende en primer lugar del tipo de institución en general de la que se trate, ya que esto conlleva en primera instancia el tipo de documentos que van a preservarse; es decir, si hablamos de una biblioteca, los materiales a preservarse son materiales documentales típicos de una biblioteca. Si hablamos de un archivo, los materiales a preservarse serán sus documentos típicos y así sucesivamente, si se trata de museos y sus acervos, o de sociedades históricas y sus

acervos, etcétera. En segunda instancia, hilando más fino, hay que pensar en particular en el tipo de institución y el tipo de materiales que ésta acostumbra guardar y pretende preservar: Esto es, si se trata de una biblioteca, de qué tipo de biblioteca hablamos: nacional, pública, estatal, regional, especializada, universitaria, escolar, etcétera, y debe tomarse en cuenta también el tipo de materiales que guarda esa biblioteca en particular. Si se trata de un archivo habría que saber a qué tipo de ellos nos referimos: histórico, gubernamental, administrativo, empresarial, etcétera, así como los tipos de documentos que guarda: registros catastrales, comerciales, académicos, médicos, policiales, de construcción, trámites, permisos, correspondencia, oficios, fichas, etcétera. Y así sucesivamente: si hablamos de un museo, de qué tipo de museo se trata, y qué tipo de colecciones quiere preservar. En este caso, no se trata tan sólo de fotografías de sus piezas, sino de muchos elementos propios e inherentes al museo, los que se vuelven parte de sus acervos a preservar: trípticos y carteles de pasadas exposiciones, programas de mano, fotografías de exposiciones, y en la actualidad, en los museos de arte contemporáneo, nuevas expresiones multimedios que son todo un reto para preservar: *performances* o instalaciones. Lo mismo vale para sociedades históricas y sus acervos documentales.

Para efectos de este trabajo utilizaremos la forma de tipificar los materiales según sus fuentes o forma de obtención, y dentro de éstas podemos hacer una segunda desagregación por tipo de material. Para la primera tipificación hay varias formas de agrupar tales fuentes, pero en términos generales se reconocen tres grandes agrupamientos de donde pueden provenir los materiales para la colección digital. En forma general son:

- 1) Materiales en soportes tradicionales que la biblioteca posee y que se han digitalizado.
- 2) Materiales originalmente digitales, ya sea de la propia institución a la que la biblioteca pertenece o adquiridos de proveedores o terceros.
- 3) Hipervínculos con otras colecciones y materiales digitales que pueden ser asociados a la colección de esa biblioteca.

Más en detalle, los tipos de materiales que son susceptibles de preservarse en sus versiones digitales por cualesquiera de la tres fuentes enunciadas son:

- Libros electrónicos, en todas sus variantes y modalidades: libros propiamente dichos, memorias, antologías, resúmenes, compendios, tesis, atlas, enciclopedias, diccionarios etcétera.
- Diarios y revistas electrónicas, en todas sus variantes y modalidades: periódicos, semanarios, *e-journals*, *e-magazines*, *blogs*, etcétera.

- Otros materiales de texto, tales como: catálogos, ensayos, reportes, listas de discusión, manuales, leyes y reglamentos, enseñanza de idiomas, etcétera.
- Imágenes digitales como: parte de colecciones de fotografías, mapas, pinturas, dibujos, diagramas, bocetos, iluminaciones, carteles, manuscritos, planos, partituras, esculturas, viñetas, patentes, estampillas, patentes, etcétera.
- Audio digital: como parte de colecciones de música, entrevistas, ponencias, mesas redondas, debates, lenguas, poesía, audiolibros, teatro, etcétera.
- Video o cine digital: como parte de colecciones de cine y video comercial o educativo; ponencias, mesas redondas, entrevistas, entretenimiento, periodismo, *instalaciones* o *performances*, etcétera.
- Bancos de datos, tablas y cartas de todo tipo: científicos, económicos, meteorológicos, geográficos, geológicos, sanitarios, estadísticos, astronómicos, jurídicos, antropológicos, educativos, etcétera. Datos espaciales y geoespaciales.
- Documentos y materiales multimedia.
- Sitios *Web* con fines de entretenimiento, de comercialización, informativos, educativos, de organizaciones, gubernamentales, de servicios, etcétera.
- Archivos documentales de gobierno, de la industria, del comercio, producidos por sus propias entidades. Esto puede incluir los correos electrónicos.
- Piezas de software, juegos digitales, realidades virtuales.

Cualquier colección de materiales de estos tipos puede ser digital de origen o es susceptible de ser digitalizada. Cuando partimos de la segunda fuente, –materiales en soportes tradicionales que la biblioteca posee y que van a ser digitalizados– es necesario, antes de tomar la decisión de procesarlos, *tamizar* cada una de esas colecciones con los criterios de evaluación para digitalización: la selección, ya que por lo general es imposible digitalizarlo *todo*. Para ello existen varios niveles de criterios según las instituciones que los establezcan: los hay de nivel bajo, medio o severo, y cada institución debe definir el nivel deseado de sus criterios para digitalizar. En términos generales, la selección de materiales para ser preservados por reformateado digital se basa en elementos tales como el valor de los documentos —físico, académico o histórico—, su uso o demanda, su condición física, la propiedad intelectual, las características del documento original y la conveniencia de las reproducciones digitales para uso y acceso, la viabilidad técnica y la no

duplicación, entre otros. Este tema es de suma importancia y debe ser siempre previo a la digitalización. Para abundar, véase el apartado 2.3 de la obra *Bibliotecas y Publicaciones Digitales*.⁵

3.2. Documentos y objetos digitales

Cuando la razón de la creación de una colección es primordialmente la preservación, o esta causa existe además de la distribución, entran en juego una serie de elementos que definen el éxito del proyecto, los recursos necesarios y el futuro de éste, y que por lo tanto deben tomarse siempre en cuenta. Al hablar de preservación documental hay algunos conceptos que resultan comunes y otros que se agrupan de modo distinto, algunos varían de autor en autor, otros ponderan más o menos algún elemento; otros más embeben este elemento en alguno superior, lo dividen o lo agrupan de manera distinta, etcétera. Independientemente de ello, es necesario definir claramente la unidad de estudio; es decir, un documento en su versión digital, ya que nosotros vamos a preservar “objetos digitales” y no simples imágenes de un documento. Esto es de importancia capital a la hora de definir proyectos.

El “Glosario de la Terminología de Archivos y Registros” –*Glossary of Archival and Records Terminology*– parte del concepto de documento electrónico al que denomina *objeto digital*: “Unidad de información que incluye las propiedades del objeto original así como también incluye métodos eficaces para realizar operaciones sobre el objeto”.⁶

De acuerdo con la Federación de Bibliotecas Digitales –*Digital Libraries Federation*– el punto de partida es su definición de *reproducción digital fiel*. Este concepto significa que el documento resultante se encuentre óptimamente formateado y reúna tres características indispensables: calidad, permanencia e interoperabilidad. [DLF, 2002]. El concepto de calidad está relacionado con su funcionalidad y su valor de uso; su persistencia tiene que ver con su capacidad de ser accesible a largo plazo y su interoperabilidad se refiere a que pueda ser accedido a través de diversas plataformas de cómputo.

Cuando a esa *reproducción digital fiel* que cumple con los elementos anteriores se le agregan los metadatos; es decir, los elementos que permiten su recuperación posterior, obtenemos un *objeto digital* u *objeto de información*. Objeto digital es por tanto, cualquier entidad documental: texto, imagen, sonido, etcétera, que haya sido codificada numéricamente bajo algún formato y ensamblada junto con algún conjunto de metadatos de tal forma que pueda

5 Juan Voutsás M., 2006, *op. cit.*, capítulo 2.3.

6 Richard Pearce-Moses, 2005, “A Glossary of Archival and Records Terminology”, Chicago: Society of American Archivists.

ser almacenada, buscada, encontrada, y usada a partir de una colección dentro de un sistema de cómputo. Estos metadatos pueden ser bibliográficos o técnicos, administrativos, descriptivos, de hipervínculo con entes exteriores al documento, etcétera. Su propósito puede querer varias cosas: identificación o recuperación, captura, origen, navegación, etcétera.

Los elementos que conforman los documentos digitales han sido en forma general identificados, pero no son agrupados de la misma manera; Por ello he tratado de establecer una resultante de lo que varios autores proponen respecto a lo que todo documento que se pretenda preservar debe cumplir y que he agrupado en mi versión en seis principios generales:

- *Selección, Calidad, Integridad, Permanencia, Accesibilidad y Funcionalidad*

Estos principios han sido establecidos al estudiar y analizar la problemática de los documentos en colecciones de bibliotecas, pero funcionan de manera muy semejante en colecciones documentales de archivos y por tanto pueden ser extrapolados en gran medida a ese tipo de ambiente, todas las proporciones guardadas.

- Principio de la *selección*: las bibliotecas digitales están conformadas por tres elementos básicos: las colecciones –obviamente digitales–, los servicios –*por supuesto en red*– y los usuarios –*naturalmente diseminados a lo largo del ciberespacio*–. De acuerdo con lo anterior una biblioteca digital les ofrece a sus usuarios servicios en red a través de sus colecciones digitales, al margen de todo lo que pueda poseer en sus acervos tradicionales. Por lo mismo es fundamental saber desarrollar una colección digital adecuadamente. Para empezar es necesario reflexionar en que al igual que con una biblioteca tradicional, no existirá una única colección documental, y que en realidad está formada por una variedad de colecciones parciales que engloban distintos tipos de materiales y formatos que conformarán lo que la biblioteca y sus usuarios conocen y agrupan como “la colección” en términos generales. De esta forma, al referirme a la “colección digital” de una biblioteca esto significará el conjunto de ellas que conforman esa colección. Por lo mismo y como sucede con toda colección bibliográfica que se respete, debe establecerse un objetivo y una política de desarrollo para la colección –o colecciones–; es decir, debe haber una *selección*. Ésta es una condición imperativa y que sin embargo es soslayada frecuentemente y se ha procedido a crear colecciones digitales de manera directa, sin ninguna

planeación. Pero su importancia es tan significativa que, sin duda, la *selección* se vuelve el primer principio de la preservación digital: saber qué se desea preservar y qué no. Cuál es el objetivo que se persigue al crear esa colección, a quién va dirigida, cuál es su nivel, alcance y cobertura, cuál su política de adquisición, permanencia y descarte de materiales, etcétera. Aquí es donde la capacitación y la experiencia del bibliotecario profesional actual es capitalizada por ser sumamente útil al momento de establecer esa política de selección de manera adecuada, la que, por supuesto, debe quedar por escrito en un documento que funcione para ese propósito como guía de largo alcance para la biblioteca. Es de suma importancia en estos objetivos y en esta selección, distinguir y precisar claramente cuáles partes de la colección son para distribuir, cuáles para preservar y cuáles para ambos propósitos.

- El principio de la *calidad* es directamente proporcional al grado en que el documento digital emule al documento original o representado y le otorgue por tanto un *valor de uso* equivalente al de ese documento. Las características que conforman la calidad incluyen conceptos como la apariencia con respecto a un original; es decir, ¿qué tanto se parece el documento digital a lo que pretende representar? Esta apariencia a su vez está definida por elementos tales como la resolución, o grado de minuciosidad de los elementos que conforman el documento, pero también el tono, color e inclusive textura; su escala –de preferencia uno a uno–, su secuencia original, su integridad, un identificador único, etcétera. Es necesario entender perfectamente este concepto de *valor de uso*: debemos estar conscientes de que un documento digital no es idéntico a un documento fijado en un soporte tradicional; puede ser muy parecido, puede ser muy semejante a su antecesor “tradicional”, pero no es idéntico. Podemos hacer que esa semejanza sea mínima, o podemos hacer un documento sumamente semejante al original; depende de técnica, esfuerzo y trabajo. ¿Qué tanto conviene hacerlo muy semejante al original? El hacerlo más y más semejante conlleva una mayor inversión de recursos, generalmente tiempo y dinero. Hacerlo poco semejante al original es más rápido y económico, pero se intuye que ello va en detrimento de la calidad del documento. Para establecer la decisión de calidad, es necesario siempre referirse al “valor de uso” del documento; esto es, a los motivos, necesidades y expectativas de los usuarios al momento de usar un cierto documento. El documento digital ¿qué tan útil será al usuario? ¿qué tanto sustituye al original? En la medida que las respuestas a estas preguntas sean más altas y el documento sea más satisfactorio y útil, diremos que tiene más calidad.

- Las posibilidades de definición de la calidad son variadas y obviamente dependen en gran medida de las expectativas del usuario.
- El principio de *integridad* es sumamente importante. Algunos autores –entre quienes me incluyo– la consideran uno de los elementos más importantes de la preservación, tanto o más que la calidad, y para el mundo de la archivística es absolutamente indispensable y fundamental. También se le denomina *autenticidad*. Este concepto se refiere al grado en que el documento digital refleja al original, pero no tanto en su apariencia física sino en su esencia, en su espíritu, en su intención. Un documento *íntegro* es el que refleja totalmente la esencia del original; es decir, no ha sido corrompido en su contexto: alterado, mutilado, interpretado, aumentado, recortado, deformado, censurado, etcétera; es confiable y por tanto aceptable. Su mensaje, autoría, fechas asociadas, lugares, etcétera, son en realidad los consignadas en el documento desde siempre; en suma: es auténtico. Aunque hubiese cambiado físicamente, en su esencia refleja de manera completa lo que el autor estableció en el documento original.
- Traduzco una definición al respecto de Luciana Duranti:

El original de un documento –esto es, la primera instanciación completa de una entidad documental que alcanzó sus propósitos– desaparece en el ambiente digital la primera vez que es salvado. Lo que nosotros recuperamos *siempre* es una copia. No podemos preservar documentos –entidades– digitales: sólo podemos preservar la capacidad de reproducirlos una y otra vez. En este contexto, una entidad digital preservada es considerada auténtica si puede considerarse que es –o aun mejor–, declararse que es una copia auténtica por el custodio quien dé fe de su identidad y de su integridad a lo largo del tiempo desde el preciso momento en que lo ingresó a su acervo y ese custodio puede documentar además correctamente el proceso de conservación –inclusive cualquier migración posterior y sus consecuencias tanto en forma como en contenido–. Ello significa que, en lo relativo a documentos en medios tradicionales, la autenticidad fue establecida siempre a través del objeto mismo, del documento, así que el custodio sólo necesitó preocuparse de que el usuario analizase el objeto y sacara sus propias conclusiones acerca de su autenticidad. Con medios digitales, lo que el usuario necesita para analizar y concluir es conocer la autoridad y la capacidad –competencia– del custodio, así como la calidad de la documentación del proceso de conservación.⁷

7 Luciana Duranti, and Thibodeau, Kenneth, 2005, “The concept of record in interactive, experiential and dynamic environments: The view of InterPARES”, in: *Archival Science*, Springer Netherlands, ISSN:1389-0166 (Print) 1573-7519 (Online), Vol. 5 Nums. 2-4, December 2005, DOI 10.1007/BF02660804. pp. 13-68.

En esta definición proveniente de la archivonomía moderna, Duranti usa la palabra “instanciación”; es decir, la representación de una abstracción a través de una instancia concreta. En la práctica sucede cada vez que tecleamos salvar ante un documento digital. Recuérdese que en el mundo bibliográfico, con las nuevas FRBR —*Functional Requirements for Bibliographic Records*⁸ a este fenómeno se le denomina *manifestación*. ¿Por qué necesitamos saber que algo es auténtico? Bien, todo investigador quiere saber si puede fiarse de sus fuentes, y cualquier lector quiere saber si mira algo verdadero o se trata de alguna falsificación, imitación, o sustituto. Además ¿no quieren los autores obtener crédito por lo que ellos crearon realmente en lugar de por alguna mala imitación? Por tanto es éste un asunto que tiene que ver con la creación, la selección y la preservación de un documento. La parte de la creación compete al autor y probablemente hasta al editor; la selección y preservación competen a la entidad custodio. Si la cadena se rompe en cualquier punto, la autenticidad se ha perdido.

- El principio de la *permanencia* tiene que ver con el concepto de que el documento estará disponible por un lapso considerable, y casi quisiéramos decir eterno. Está asociado con su presencia, su seguridad, y por supuesto, con la duración y continuidad de su soporte. Atención, muchas personas tienden a confundir este principio con el de *accesibilidad*, el cual tiene que ver con que el documento, aun si existe, pueda ser accesible. Pero son dos conceptos distintos, trataré de explicar brevemente esta diferencia. Para ello es necesario que podamos distinguir bien entre *almacenamiento permanente seguro* —la “permanencia”— y el *acceso futuro* —la *accesibilidad*—. Para asegurarnos el almacenamiento permanente seguro de los objetos digitales, son necesarios procedimientos y técnicas adecuadas para mantener archivos documentales a largo plazo. Tales archivos deben preservar la *cadena de bits* y su formato más sus respectivos metadatos que aseguren la descripción y la búsqueda. Se deben diseñar por tanto y llevar a cabo esas técnicas y procedimientos meticulosos para conservar los soportes documentales y sus contenidos digitales.

El acceso futuro de los objetos digitales almacenados, la accesibilidad es otra cuestión. ¿Cómo podemos estar seguros de que —habiéndose conservándose en buen estado— vamos a poder ver o ejecutar estos archivos dentro de veinte, cincuenta o doscientos años? Es decir: ¿cómo

8 FRBR. 2007, “Functional Requirements for Bibliographic Records - Final Report”, IFLANET Publications, disponible agosto 2008 en: http://www.ifla.org/vii/s13/frbr/frbr_current_toc.htm.

garantizaremos nuestra capacidad de reproducir correctamente las *cadena de bits*?

- Como parte del principio de la *accesibilidad*, tenemos el concepto de *interoperabilidad*. Este último significa que el documento será accesible a lo largo de variadas plataformas y programas de computadora. Esto es, el documento no debe estar atado a formatos específicos de un proveedor de *software*, o a una marca de computadora o de cámara fotográfica, etcétera. Un ejemplo muy ilustrativo lo vemos cuando intentamos abrir un documento en una página *Web* y nos aparece el mensaje “este documento sólo puede verse en *explorer* versión tal...”. Ese documento no es interoperable. Un documento en verdadero formato html es visible en prácticamente cualquier navegador: *explorer*, *netscape*, *firefox*, *ópera*, *chrome*, *safari*, etcétera; incluso a lo largo de varias versiones de cada uno de ellos. El secreto de la interoperabilidad es preferir siempre el uso de “estándares abiertos” y evitar los “estándares propietarios”, ya que a veces, en aras de una cierta comodidad y/o rapidez de creación, nos atamos a una plataforma específica de equipo o programas. En suma, en la medida que un documento es más interoperable durante su creación reduce el problema de la *accesibilidad* futura y obviamente lo opuesto puede afirmarse en sentido inverso. Los proyectos de preservación y distribución digital que se han apegado a los estándares de interoperabilidad han demostrado una y otra vez que el apegarse a estándares paga a la larga, y reduce las inversiones y reconversiones futuras de esos proyectos. Y lo opuesto se ha demostrado en los proyectos que no lo hicieron así. Lo demás de la accesibilidad tiene que ver con aspectos como: formatos y plataformas tecnológicas, emulación, migración, etcétera.
- El último principio mencionado en la preservación es el de la *funcionalidad*. Más formalmente, la funcionalidad consiste en las características interconstruidas en una interface de búsqueda que determinan la facilidad con que los usuarios pueden formular búsquedas y obtener resultados.⁹ Tratemos de explicar este concepto: su trasfondo está asociado a la facilidad con que los usuarios puedan acceder a un documento y a sus documentos relacionados. Pero esa facilidad depende de las capacidades que puedan lograr las interfaces de búsqueda de documentos, las que a su vez dependen de las características de información que puedan obtener del interior de los propios documentos. En conjunto la funcionalidad se mide por la suma de las capacidades que contengan

los documentos, programas, sistemas, etcétera; si observamos que una parte muy importante de esta funcionalidad proviene del propio documento, entenderemos la importancia de que un documento digital tenga funcionalidad intrínseca; de otra forma afectará negativamente al resto de los componentes informáticos que lo accederán.

De ahí la importancia de los metadatos que acompañan a un “objeto digital”. Sin metadatos el documento digital no podrá ser accedido en forma alguna; no tendrá funcionalidad en lo absoluto. Algunos metadatos sencillos incrementarán en cierta forma esa funcionalidad en la medida que el documento podrá ser recuperado y asociado con otros. Buenos metadatos nos aproximarán a una buena funcionalidad, como el grupo de metadatos “típico” o “estándar”: metadatos descriptivos, técnicos, estructurales y legales. Obviamente deseamos llegar algún día no tan sólo a una buena, sino a una óptima funcionalidad para que el documento tenga óptima recuperación y acceso.

Para poder entender más a fondo este concepto de la “óptima funcionalidad” puede acudirse a los conceptos de la *Biblioteca Semántica* y la *Web Semántica*. La *Biblioteca Semántica* es la tendencia evolutiva de las bibliotecas; las bibliotecas digitales se van convirtiendo poco a poco en *bibliotecas semánticas* y todo parece indicar que ése será su siguiente estadio. Este concepto se basa a su vez en el principio de la *Web Semántica*. En una explicación sencilla, la *Web Semántica* es una *Web* extendida, dotada de mayor significado y en la que cualquier usuario de Internet podrá encontrar respuestas a sus preguntas de forma más rápida y sencilla gracias a una información mejor definida. Las páginas *Web* hasta ahora están hechas para ser leídas por seres humanos, no por máquinas, pero aquello que busca dentro de esas páginas para hallar lo deseado son máquinas, no seres humanos. Al dotar a la *Web* de más semántica para sí misma, tendrá por lo tanto más significado intrínseco para sus programas y buscadores, y gracias a ello se podrán obtener soluciones a problemas habituales en la búsqueda de información gracias a la utilización de una infraestructura común, mediante la cual sea posible compartir, procesar y transferir información de forma sencilla. Esta *Web* extendida y basada en significado no tan sólo humano sino también computacional, se apoya en lenguajes universales que resuelven los problemas ocasionados por una actual *Web* carente de semántica en la que, en ocasiones, el acceso a la información se convierte en una tarea difícil y frustrante.

Una *biblioteca semántica* por tanto será una biblioteca digital cuyas colecciones estén formadas por *documentos funcionales*. Un documento

funcional comienza con un objeto digital, como sucede en la actualidad en una colección de biblioteca digital típica, documento que como ya mencionamos consta de una “reproducción digital fiel”, que incluye sus características de selección, calidad, permanencia, integridad, etcétera, pero que no se queda en sólo algunos simples metadatos agregados para identificarlo. En resumen, la funcionalidad de un documento está dada por características inherentes de calidad aunadas a adecuados metadatos, sobre todo los de hipervínculo los cuales lo relacionan con otros documentos semejantes —es decir, funcionales— de manera tal que el conjunto documental así relacionado tendrá un significado procesable por computadora y esto permitirá que los sistemas puedan darle a los documentos prelación, pertinencia, pertenencia, jerarquía, relación, etcétera, y sea posible lograr así una búsqueda más inteligente y por tanto una recuperación más precisa y útil para el usuario. Esto debe incrementar sensiblemente la reutilización y el hallazgo de esos documentos, la adecuada minería de datos y la integración, todo bajo un esquema de interoperabilidad entre sistemas. Como conclusión de este último punto dentro de la preservación digital, cabe decir que no tan sólo la información intrínseca del documento es importante ésta está relacionada con otros elementos que forman un conjunto que permite crear efectivamente un objeto completo y de calidad que cumple con toda una serie de requisitos que lo harán susceptible de ser bien preservado. Por tanto todo documento que se pretenda preservar deberá cumplir lo más posible con los seis principios de los objetos digitales que exige la preservación. En la medida en que cumpla con ellos, ese documento será un buen —u óptimo— objeto digital.

Como resumen, un *buen objeto digital*:

- 1) Debe provenir de una política específica de selección, y por tanto cumplir con las prioridades y objetivos de la colección.
- 2) Tendrá calidad intrínseca: apariencia adecuada, resolución, tono escala, secuencia, integridad, compresión sin pérdida, identificador único, etcétera.
- 3) Será auténtico: de alguna forma garantizará que aunque su estructura física cambie, su contenido, su “esencia” será la misma; esto es, seguirá siendo íntegro, auténtico.
- 4) Será permanente; esto es, deberá existir la intención y responsabilidad de una persona o institución identificados, de que ese objeto permanecerá siendo accesible a lo largo del tiempo. Estará representado en

un formato digital que soporte el uso actual para el que fue diseñado así como usos futuros, y será susceptible de ser copiado a otros formatos para ese propósito.

- 5) Tendrá accesibilidad; esto es, la institución que lo custodia tendrá acceso a las tecnologías que permiten acceder a ese documento, o bien el objeto digital habrá sido modificado, actualizado, emulado, etcétera a las nuevas tecnologías al alcance de la institución. Tendrá por tanto interoperabilidad —será intercambiable entre plataformas—, habrá sido hecho bajo algún estándar o buena práctica reconocidos y sólo se alejará de ellos debido a razones claras y bien documentadas.
- 6) Será funcional; esto es, tendrá una riqueza notable de metadatos, tanto descriptivos, técnicos, estructurales, legales y de hipervínculo, de tal forma que será semánticamente rico, y por tanto estará ligado de tal forma con otros documentos funcionales de manera tal que permitirá su óptimo reuso, búsqueda, integración, minería de datos e interoperabilidad y proveerá programas, sistemas e interfaces de una lista completa de sus contenidos.

Como ha podido observarse con estos principios es sumamente importante que toda biblioteca digital establezca con precisión sus objetivos para desarrollar sus colecciones. Las bibliotecas deben crear buenos objetos digitales para conformarlas, pero el concepto de lo que significa un buen objeto digital depende en mucho de para qué se crean esas colecciones y sus objetos. Por tanto la primera decisión a tomar, como ya hemos mencionado, será establecer si vamos a crear objetos digitales para su preservación, su distribución, o ambas cosas, ya que como hemos empezado a establecer, las características y los costos de cada uno de estos objetos son sumamente distintos. Nunca habremos de emprender proyectos de preservación con especificaciones pensadas para la distribución ni viceversa. Las consecuencias pueden ser lamentables.

Si partimos de la premisa de que *digitalizar* un documento es un proceso de representación arbitraria del mismo por medio de un cierto conjunto o secuencia de números o dígitos. Hay más de una representación posible para cada documento; es decir, pueden establecerse diversas maneras o “patrones” para digitalizar un documento. A un tipo de representación dada bajo una cierta convención arbitraria preestablecida se le conoce como un “formato”. Hay tantos formatos posibles como convenciones numéricas de representación puedan establecerse para un documento —es decir, infinitas—. Muchas de estas convenciones escogidas tienen que ver con ciertas características deseables para ese tipo de documentos. Pero no todos los formatos

tienen las características ideales para representar adecuadamente a un documento; esto más bien depende de las necesidades y propósitos del que lo crea. Por ello un formato dado es una representación de un documento que tendrá ciertas características que deseamos en el mismo. No profundizaré más acerca de los formatos, el tema ha sido tratado ampliamente en la obra *Bibliotecas y Publicaciones Digitales*.¹⁰ Mencionaré sólo los elementos indispensables para continuar con la ilación de esta obra.

Hemos mencionado anteriormente el concepto de “formateado óptimo” de un documento como parte de su calidad. Existen formatos para representar múltiples tipos de documentos: para texto, imágenes, audio, video, etcétera. Para cada uno de estos tipos de documentos existe una variedad enorme de formatos creados por los fabricantes de equipo y programas; instituciones, asociaciones, etcétera. ¿Cuáles son los aspectos más importantes sobre los que un bibliotecario debe decidir respecto a formatos para que éste sea óptimo?

- 1) *Calidad de almacenamiento del formato*. Uno de los elementos que definen la calidad de un documento es su “resolución”; esto es, el número de elementos que se consignan por unidad de superficie. La resolución se expresa en distintas unidades en función del tipo de documento. Esto es, un documento “imagen” o “video” por lo general se expresa en puntos por pulgada –DPI– o Megapíxeles –millones de puntos que componen la imagen–. Un audio se expresa por lo general en kbps *kilobytes por segundo*. En todos los tipos de documentos, un documento con mayor resolución tiene mayor calidad, ya que cuenta con más datos para las bocinas, pantalla, impresora o dispositivo donde vaya a ser reproducido. Por lo tanto desde el punto de vista de la calidad se desea una alta resolución. Por otro lado, a mayor resolución de un documento, su tamaño se incrementa exponencialmente, ya que se ocupa más espacio para guardarlo y más tiempo en transmitirlo, lo cual incrementa los costos de guardarlo y transmitirlo. Así, tamaño y costo de un documento son factores totalmente entrelazados y directamente proporcionales.

En todos los proyectos de colecciones digitales aparece siempre esta interrogante para el bibliotecario: el balance adecuado entre calidad y consumo de espacio. A mayor calidad del documento digitalizado crece el espacio requerido para guardar el material en una forma no-lineal; y mayor espacio implica mayor costo de almacenamiento y mayor tiempo

de transmisión del documento. De manera inversa, menor calidad del documento implica menor espacio de almacenamiento y de transmisión; es un documento “económico”. Pero no tan sólo eso, es un documento práctico para el usuario, ya que no requiere tampoco de grandes cantidades de tiempo en descargarlo ni de espacio al poder ser almacenado incluso en un dispositivo portátil. Los formatos responden en gran medida entonces a esta necesidad de calidad contra economía. Hay formatos que le dan preferencia a la calidad y otros que se la dan a la economía. Esto marca una separación definitiva, por lo menos por la próxima década, de los formatos que se utilizarán para preservar y distribuir material por parte de las bibliotecas. La confusión respecto a las características y costos de unos y otros marca muchas veces la diferencia entre el hecho de que un proyecto de digitalización se realice bien o no; por ello es necesario que los bibliotecarios estén conscientes de esta característica en un documento digital, aunque sea de manera general, así como de sus formatos y costos asociados.

- 2) *Vigencia y estandarización del formato.* Los formatos, como muchas otras entidades del mundo del cómputo, aparecen y obtienen un mayor o menor grado de aceptación por parte de los usuarios y por tanto tienen una vigencia durante algún tiempo, después del cual tienden a ser menos usados y/o a desaparecer. Por ello es importante estar conscientes de cuál es el periodo de vigencia del formato ¿es un formato “joven”, con vigencia por delante o es un formato que ya va de salida? Hay que evitar tanto estos últimos como el último “grito de la moda”. Los formatos maduros, vigentes y en su apogeo son siempre el punto medio ideal. De la misma forma es muy importante conocer el grado de estandarización del formato ¿quién lo ha producido; un fabricante o una institución? ¿es un estándar *de iure* o *de facto*? ¿es sólo un formato propietario de un fabricante? Los bibliotecarios siempre han estado muy conscientes de la importancia de los estándares; este caso no es la excepción. Deben preferirse siempre los formatos estándar, recomendados por otras instituciones de importancia y evitarse los formatos propietarios.
- 3) *Versiones de los formatos.* Al igual que muchos otros elementos técnicos, los formatos tienden en la actualidad a ser actualizados de tiempo en tiempo con nuevas “versiones” por parte de la institución que los produce. Es muy importante que si usamos un formato en nuestras colecciones, estemos atentos a las características de esas nuevas versiones para ver si éstas son de nuestra conveniencia. Habrá que decidir si los nuevos documentos se registran con la nueva versión del formato o

con alguna de las anteriores. Para ello es de suma importancia verificar la compatibilidad de los documentos bajo la nueva versión con los documentos en versiones anteriores, y evitar de esta forma inconsistencias entre documentos de la misma colección.

Finalmente, para terminar este apartado, es importante hacer una reflexión sobre los criterios para digitalizar una colección que se encuentra en soportes “tradicionales”. Este tema fue tratado en el capítulo 2.3 de la obra ya citada y ahí puede abundarse si se desea más información.¹¹

Dado que las colecciones digitalizadas facilitan en mayor o menor grado el cumplimiento de sus dos propósitos básicos: preservar o distribuir, es imprescindible establecer previa y claramente criterios para decidir si se va a realizar esta tarea. Con frecuencia nos pasa por la mente *mejor digitalizar todo*. Pero si bien la idea es atractiva de inicio, presenta obstáculos enormes para ser llevada a cabo. El número de documentos producidos desde el advenimiento de la imprenta acumula ya muchos millones de ellos. Aun sin pensar en el todo, cualquier biblioteca con un acervo importante tendría enfrente la tarea de convertir millones y millones de páginas para lo cual requeriría de enormes recursos y enfrentaría obstáculos de todo tipo antes de poder ser llevada a cabo. Conscientes entonces de que no es recomendable digitalizar en todos los casos toda nuestra colección, sino sólo algunas partes de ella, es importante establecer *perfectamente* cuáles deben ser los criterios para que una organización, entre ellas la biblioteca, emprenda un proyecto a este respecto. La pregunta entonces es ¿qué debe digitalizarse, y con qué prioridades? ¿cómo debe hacerse? ¿por qué queremos hacerlo? Para contestar estas preguntas conviene revisar lo que ya se ha establecido acerca de ellas en algunas de las bibliotecas que ya han emprendido este tipo de tareas y que se consideran *buenas prácticas*. En términos generales, los criterios para la selección de materiales para ser preservados por reformateado digital se basan en el valor, uso, condición, características del documento original y conveniencia de las reproducciones digitales para su uso y acceso. Bajo estos rubros, los criterios pueden ser más o menos detallados y estrictos y dependen de cada biblioteca u organización. Los criterios aprobados por una biblioteca deben aplicarse sin excepción a todas y cada una de las partes de la colección susceptibles de ser digitalizadas, con objeto de evitar errores, inconsistencias y proyectos sin sustento de largo plazo. Esta definición de criterios permitirá además clarificar y establecer a la vez aspectos técnicos como son los formatos que se emplearán en la digitalización, lo que dependerá de si sus fines son

la preservación y/o la distribución, y determinará también los tipos de compresión permisible, la resolución; los aspectos de tipo legal; el “uso ético”, etcétera.

3.3 El problema de los medios de soporte

Uno de los debates de mayor intensidad en los últimos años ha sido el relacionado con la duración de los medios utilizados para darle soporte a la información digital. Como mencioné previamente hay una enorme preocupación acerca de la duración de esos soportes actuales, los cuales, en términos generales, no son más duraderos que los anteriores a la electrónica. Al contrario; tienden a durar menos, o al menos así lo aparentan.

Es muy conocida la gráfica de Paul Conway, acerca de “la ironía de los medios modernos” en la cual se grafica la expectativa de vida de los medios frente a su densidad de grabación. Ahí puede verse claramente que a lo largo de los siglos, los medios escriptorios han aumentado sensiblemente su densidad de grabación, pero ha disminuido de igual forma su expectativa de vida [Conway, 1996].

Esto debe reflexionarse con cuidado. En efecto, algunos de los soportes de la era electrónica son inútiles después de algunos años, como las cintas magnetofónicas, los casetes y los disquetes. Después de cierto tiempo el plástico se pone quebradizo y el óxido férrico cae de su superficie, lo cual inutiliza la información contenida en ellos. Pero éste no es el caso general de los medios modernos sino sólo de algunos de ellos. Al hablar de soportes, de hecho el problema con algunos plásticos es que son eternos, o casi. Plásticos tan cotidianos como el polietileno, el poliéster, etcétera, son enormemente duraderos. A los envases de polietileno o policarbonatos deben agregárseles por ley en muchos países sustancias adicionales puestas a propósito para lograr la degradación de las moléculas. De otra forma, al irse a la basura durarían 10,000 años, aun enterrándolos.

No sucede lo mismo con el recubrimiento del soporte, la capa donde realmente están almacenados los bits. La más duradera es aquella donde los bits están moldeados sobre la misma superficie, como es el caso de los CD y los DVD grabados en fábrica. Los otros recubrimientos no son tan duraderos, sean ferromagnéticos, como en disquetes, cintas magnéticas, casetes, etcétera, o se trate de capas de laca fotopolimerizable adicionada, como en CD-R y DVD-R, CD-RW y DVD-RW, DVD+R, o DVD+RW. Según el tipo de disco y sus condiciones de almacenamiento, éstos podrían durar de 1 a 75 años. Lo ideal es almacenarlos entre 10° y 15° celsius, con un 25% a 30% de humedad relativa, sin luz UV y con atmósfera limpia de gases y partículas.

Duración de soportes de almacenamiento en función de humedad y temperatura

Humedad Rel.	25 % H.R.	30 % H.R.	40 % H.R.	50 % H.R.	50 % H.R.
Dispositivo Temperatura	10° C	15° C	20° C	25° C	28° C
Cinta Magnética D3	50 años	25 años	15 años	3 años	1 año
Cinta/Cartucho Magnético	75 años	40 años	15 años	3 años	1 año
Cd-ROM / DVD	75 años	40 años	20 años	10 años	2 años
Cd-R	30 años	15 años	3 años	9 meses	3 meses

Tomada de: Digital Preservation Coalition, disponible enero 2007 en: <http://www.dpconline.org-graphics/medfor/media.html>.

Es cierto entonces que los soportes de los registros digitales actuales no duran muchos años. Pero es que *no están hechos para durar*. Están hechos para ser sustituidos en pocos años por nuevos soportes de mucho mayor capacidad. Y por tanto no es necesario que duren. El verdadero problema de los registros documentales es que la tecnología usada para leerlos no haya sido actualizada y no se encuentre disponible; el verdadero problema no es la permanencia de los soportes sino la accesibilidad futura de la información.

Muchos autores en la bibliotecología moderna cuestionan enormemente el hecho de convertir las colecciones para digitalizarlas. Se argumentan muchas razones: que la tecnología todavía no está madura para ello, que los soportes duran poco, que los documentos no son iguales al original, que la tecnología se hace obsoleta; que hay que reconvertir y eso cuesta, etcétera. Todas estas razones son reales y por tanto innegables, pero debemos estar conscientes de una realidad inmensa e inamovible: no existe el medio ideal —no ha existido y no parece que existirá todavía en un futuro cercano—. Sé que ésta es una afirmación delicada y debemos por tanto analizarla más a fondo.

Deanna Marcum, bibliotecaria de servicios de la Biblioteca del Congreso de los E.U.A, menciona una historia en la que un senador del reino ostrogodo en la Italia del siglo VI llamado Magno Aurelio Casiodoro escribió en sus “cartas” acerca de la lejana invención del papiro con toda elocuencia: Marcum introdujo esta cita para quejarse después de la fragilidad del papiro y llega a la siguiente conclusión:

...sólo una fracción de lo que el mundo antiguo consignó en papiro ha llegado hasta nosotros. El calor y la humedad del Nilo han tenido que ver en ello, junto con vandalismo, incendios y otras catástrofes naturales y humanas, las cuales borraron numerosos registros en papiro de los antiguos imperios de Persia, Grecia y Roma.

Del papiro Marcum pasa a la fragilidad de los registros digitales. Así que ni papiro ni digital son buenos soportes. Cabe entonces que hagamos la siguiente

pregunta: ¿existe un soporte *ideal* para los registros documentales? La piedra y la arcilla cocida han probado ser lo más duradero. Pero la piedra no era fácil de producir ni de almacenar; la arcilla cocida era fácil y barata de producir, mas no de almacenar. Ambos medios ocupan mucho espacio y no tienen una gran densidad de grabación. Sólo sirven para textos; muy poco para imágenes y nada para audio o video. El papiro, pergamino o papel son mucho más ligeros y flexibles, pero son muy susceptibles a la humedad y al fuego; no duran ni con mucho lo que la piedra. Adelantemos un poco la respuesta la cual es –obviamente– que no existe un medio ideal como soporte para documentos; como casi todo en esta vida, cada uno tiene ventajas y desventajas, mismas que deben ser ponderadas de acuerdo con cada situación.

A mediados del siglo decimonono el papel hecho de celulosa comenzó a ser fabricado en grandes cantidades con lo que la demanda pudo por fin ser satisfecha con creces. El papel era además –gracias a la industria– más uniforme, blanco, delgado pero resistente, abundante y, sobre todo, barato. Este tipo de papel sí desplazó en unos cuantos años al papel de trapo, con el cual por lo mismo prácticamente no coexistió. Maravillosa como puede parecer esta técnica, incluía un grave problema; el papel producido así en esa época era ácido, lo que a la larga lo vuelve quebradizo y lo destruye. El papel de trapo dura siglos; el papel de celulosa ácido unas pocas décadas. Entre más ácido dura menos; y es por lo general el papel más barato. La química del papel de celulosa no era bien comprendida en ese entonces, ni tampoco varias décadas después de su creación.

En 1897, John Russell Young, encargado de la Biblioteca del Congreso de los E.U.A. resaltó el creciente problema del rápido deterioro del papel de los libros recientes, e hizo hincapié en los problemas derivados del proceso del mismo y las preocupantes consecuencias.¹² Así que presentó un plan al congreso en el cual proponía que los materiales entregados a la biblioteca por depósito legal deberían estar impresos en papel de buena calidad. Él escribió:

Sabemos que el papel barato debe existir; es uno de los desarrollos necesarios para los negocios y para la popularización de los libros; pero ello no es razón para que las bibliotecas no deban ser protegidas de ello. El costo de unas pocas copias de cualquier publicación en papel de alta calidad sólo significaría unos pocos centavos extras y unos momentos de demora más en las prensas. Por ello propongo que las publicaciones para el depósito legal deban ser exigidas obligatoriamente

12 John Russell Young, 1897, *Report of the Librarian of Congress*, December 9, 1897 (Washington: GPO, 1897), pp. 49-50.

impresas en un papel que cumpla un mínimo de calidad y durabilidad. Tal medida asegurará la permanencia por mucho tiempo más de las colecciones en ésta y en otras bibliotecas semejantes.¹³

El plan de John Russell no fue implementado, ni ningún otro alternativo. Nadie más pareció darse cuenta de la gravedad del problema, o al menos nadie más le dio importancia hasta bien entrado el siglo XX, y aún así pasarían otros cincuenta años antes de que comenzara a fabricarse el papel con acidez neutra, y otros treinta y siete años más para que hubiera una norma norteamericana para papel con calidad de preservación.

¿Qué tan grave es el problema de los *libros quebradizos*? Para darnos una idea, en 1986 un comité de preservación documental hizo en los Estados Unidos un estudio al respecto el cual, en resumen dice que el papel de la mayoría de los libros publicados en todo el mundo en los últimos 125 años fue de este tipo, y que por tanto su periodo de vida será corto. De hecho, muchos de ellos ya han empezado a desbaratarse. Se calcula que solamente en ese país diez millones de volúmenes desaparecerán por esta causa en los próximos veinte años, [Committee..., 1986] mucho antes que los libros de los siglos XVI, XVII o XVIII, los cuales, al haber sido impresos en papel de trapo, durarán varias centurias más.

Desde la década de los cincuenta, en cuanto algunas bibliotecas se dieron cuenta de este problema, se desarrollaron procedimientos para “desacidificar” el papel, lo cual lograron con razonable éxito. El problema es que es un proceso caro, minucioso que requiere de mucho tiempo y paciencia, y que por lo mismo es impráctico para colecciones de un tamaño considerable o masivo. Por lo mismo la única salida práctica al problema es copiar esos materiales en nuevos soportes, aun si se toma en cuenta el costo y tiempo de esta operación, ya que de otra forma en unos años más se perderán gradualmente. En la unión americana, después de que este problema se empezó a conocer se hicieron dos cosas: se votaron disposiciones para fabricar papel más duradero que permitiera mayor preservación, y la comisión persuadió al congreso americano para que estableciera y patrocinara un programa de preservación nacional de *libros quebradizos* microfilmándolos.

Respecto a lo establecido por el Comité de Preservación de la Biblioteca del Congreso, se creó el programa de preservación documental por medio de la microfilmación. Hay que recordar que la digitalización documental, tal como la conocemos ahora, no existía. Este programa comenzó en 1989 y fue administrado por la “Oficina de Preservación del Fondo Nacional para las

Humanidades” –*Office of Preservation of the National Endowment for the Humanities*–, una agencia gubernamental que les provee fondos a las bibliotecas para microfilmarse sus acervos. Se proyecta que en un par de años más, cuando termine, habrán microfilmado poco más de un millón de volúmenes en 20 años.

Toda esta historia se ha traído a colación para que nos demos cuenta realmente de que la historia del texto impreso en soportes flexibles como el papel es una historia de total éxito si la vemos como vehículo para la cultura, pero es una historia llena de altibajos desde el punto de vista de la preservación y presenta grandes problemas para el futuro, especialmente para las instituciones que tienen la responsabilidad de la preservación documental a largo plazo. Partiendo de este análisis no puede afirmarse de una manera categórica que el papel sea el medio ideal para preservar, con lo que se da respuesta a la pregunta que nos hacíamos al principio de este apartado. Como hemos podido analizar aun el mismo papel –el material escriptorio por excelencia–, tiene sus deficiencias y sus riesgos. Reitero, no podemos negar lo que el papel y el libro han significado como vehículo para la cultura mundial. Muy poco sería este mundo y su cultura sin ellos y en modo alguno quiero referirme al papel o al libro de forma peyorativa; sólo quiero recalcar el hecho de que, desde el punto de vista de la preservación, los textos en materiales flexibles, finalmente el papel, también presentan una serie de inconvenientes serios y no están en modo alguno exentos de riesgos con los que los responsables de la preservación deben contender.

A pesar de estos riesgos no podemos cancelar de golpe la cultura del papel. Seguimos creyendo en él; seguimos confiando en él; pero es muy importante que estemos conscientes de que no es la única ni la ideal solución al problema de la edición y preservación y que –con todos los inconvenientes y riesgos que ello implica– *debemos* empezar a considerar otros soportes documentales. Nuevamente hay que hacer coexistir al papel con otros soportes y dejar que el tiempo y el usuario establezcan su preferencia. El punto es hacer notar que el papel tiene ventajas distintas a las de un soporte digital desde el punto de vista de la preservación, pero que no es el medio ideal para ello y nuestra búsqueda debe continuar. Por más que amemos al papel no podemos quedarnos estacionados en él. No porque sea obsoleto, no porque queramos olvidarlo, sino porque es necesario movernos para continuar nuestra jornada como seres humanos en el mundo de la cultura y de la edición preservables. Y si no es por estas razones, no olvidemos que no podremos sostener la producción y demanda de papel de la actualidad por otros veinte años sin causar graves daños ecológicos al planeta. Ya los hemos causado.

El análisis no termina aquí. Hemos hablado del soporte *papel*, pero no ha sido el único. Hagamos un rápido análisis de los soportes “tradicionales”

—previos a la era digital— para otros tipos de documentos. Si nos referimos a las imágenes —no impresas en papel, puesto que ese soporte ya lo analizamos— tenemos, por ejemplo, una gran cantidad de negativos en vidrio y en nitrocelulosa, así como transparencias creadas a lo largo de siglo y medio de fotografía. Sobre todo los primeros, los de vidrio y de las primeras épocas del rollo flexible, son muy delicados y frágiles. El soporte de vidrio no se degrada, pero su recubrimiento fotográfico sí: colodión, albúmina, etcétera. Algo semejante pasa con los negativos y transparencias de celuloide de la primera mitad del siglo XX; sólo que en este caso el soporte es el que corre riesgo. Y al igual que con el papel, el agua y el fuego los destruyen.

Si continuamos con los registros de audio, el problema es semejante. Los primeros registros sonoros estaban en cilindros de cartón recubierto con una película metálica. Por supuesto no era el medio ideal; rápidamente fue reemplazado por los discos de acetato, los cuales son duraderos mientras se almacenen, pero extremadamente frágiles y quebradizos al uso. Como muchos hemos observado, un disco de 78 r.p.m. hecho de acetato se quiebra y despostilla al menor golpe. Los discos de vinilo mejoraron el panorama en cuanto a flexibilidad, resistencia y duración, pero llegó un momento en que eran demasiado voluminosos para tan sólo contener 45 minutos de grabación, y de todos modos eran susceptibles de rayarse. Ya hemos mencionado los problemas de las cintas magnetofónica y los casetes, los cuales, con fines de preservación tampoco duran mucho.

Finalmente podemos notar que con los registros de cine y video pasa algo parecido. Los registros de las primeras décadas del cine estaban hechos en soportes de muy poca duración, como celuloideos de baja calidad o algo aún peor: sobre soportes de altísima flamabilidad como la nitrocelulosa —material que en ciertas condiciones, arde espontáneamente. De este material, muy popular en su momento, lo que no ha sido copiado hoy en día a otros soportes prácticamente ya no existe. Si bien el material de la segunda mitad del siglo XX es de mayor calidad, no es un soporte ideal y tiene sus desventajas: es susceptible al fuego, al agua, y al desgaste mecánico por uso. Con el video analógico pasa lo que con las cintas magnetofónicas: ni el soporte plástico ni el óxido férrico son duraderos. El video de las primeras décadas debe ser copiado o se destruirá en breve, y el más reciente no durará muchas décadas.

Estas últimas reflexiones nos permiten llegar a la conclusión central de este apartado. No podemos oponernos a los soportes digitales con el argumento de que son riesgosos, inciertos, perecederos y frágiles. Desde que abandonamos las tabletas de arcilla para acá, ésta ha sido precisamente la historia de los documentos y la problemática de la preservación. Llevamos tres milenios conteniendo con esos riesgos y con esa incertidumbre, y hemos visto que

con ellos han coexistido durante mucho tiempo variados soportes. Debemos analizar las ventajas y desventajas de los medios digitales desde una perspectiva más objetiva frente a sus equivalentes “tradicionales” para saber en qué casos, bajo qué condiciones, con qué costos y con cuáles reglas debemos internarnos en esos procedimientos. Pero no podemos esperar garantías ni seguridad absoluta porque no las ha habido nunca. Sabemos de antemano que los electrónicos no son los soportes ideales, pero espero que haya quedado establecido que nunca los ha habido. Por tanto debemos a toda costa estudiar a fondo el proceso de la preservación digital para conocer a fondo todas las variables que intervienen en el proceso. Debemos evitar cometer errores y hacerlo bien desde la primera vez, pero en todo caso, debe quedar muy claro que el éxito o el fracaso no serán culpa del medio, sino de nuestras decisiones y acciones.

3.4 Los medios y la obsolescencia tecnológica

Cuando el custodio de los bienes documentales —bibliotecario o archivista— no hace planes para una preservación de largo plazo de los materiales en un cambiante medio tecnológico y suceden los “accidentes” de pérdida de materiales, por lo general *se mata al mensajero por traer malas noticias*. En este caso —y como ha pasado muchas veces— se tiende a señalar al soporte o a la tecnología como el culpable por los malos resultados obtenidos. Esto también sucede cuando los planes han sido correctamente elaborados pero con el tiempo nadie se acuerda de seguirlos. Pero si lo reflexionamos un poco más a fondo, observaremos que el soporte tecnológico no es culpable, lo es sólo el mensajero.

Como ya hemos mencionado, por algún tiempo la preocupación giró acerca de la duración de los soportes de los archivos electrónicos: ¿cuánto dura una cinta, un disquete, un cd-rom? etcétera, ya que el pensamiento era guiado linealmente respecto a la preservación basada en la duración de los soportes en papel. Durante décadas hubo docenas, cientos de estudios reportados en donde se *envejeció* artificialmente a los discos, cintas, etcétera, con el fin de poder pronosticar su duración final verdadera. En los últimos años se ha empezado a pensar en los soportes de manera más independiente que en los anteriores, y a enfocarse cada vez más el verdadero problema de los soportes tecnológicos actuales: la obsolescencia tecnológica. De hecho, ya hay proyectos alrededor de la idea de los “bits sin ácido” —*acid free bits*— donde la idea central, al igual que la desacidificación del papel, es quitar de los registros electrónicos, en lo posible, los elementos que pueden reducir su accesibilidad en “un” futuro. Aunque el nombre así lo sugiere, nada absolutamente en esta

idea tiene que ver con la duración de los soportes, sino con el problema de la accesibilidad de la información, lo cual se deriva de la vigencia tecnológica.

En efecto, la obsolescencia tecnológica representa una amenaza mucho mayor para la información electrónica que la inherente fragilidad física de sus soportes [Mallinson, 1986] y [Gavrel, 1986]. Al no haber hoy en día una solución universal contra este problema, se ha establecido el consenso de utilizar cuatro técnicas que tienden a optimizar la preservación digital y que resuelven tanto el deterioro de los soportes como la obsolescencia tecnológica. Pero ninguna es total, cada una resuelve una parte del problema; no obstante, usadas en conjunto han probado ser una solución suficientemente buena por el momento. Decía Teodoro Roosevelt que en un momento de decisión, lo mejor que puedas hacer en ese momento siempre será la mejor decisión; la peor siempre será no hacer nada. Las cuatro técnicas mencionadas, en resumen, son:

Réplica, Recopia, Migración y Emulación

- La “réplica” consiste en crear y mantener varias copias de la información en distintos lugares. Si sólo existe una copia de la información electrónica documental, en caso de falla del computador, daño o accidente debido a fuego, inundación, terremoto, etcétera, la información puede verse seriamente comprometida. El hecho de crear y guardar varias copias de la información y guardarlas en lugares distintos evita este riesgo y multiplica las probabilidades de que cierta información sobreviva al tiempo y a los percances posibles.
- La “recopia”, también llamada “renovación”, “rejuvenecimiento” o “refrescado” —*refreshing*— consiste en la sencilla técnica de copiar los registros electrónicos, de cuando en cuando, hacia otros soportes más nuevos, más “frescos” y de mayor capacidad. En esta técnica se copia la imagen del archivo tal cual está, sin modificación alguna. Se entiende por ello mismo que los formatos internos de los documentos no cambian ni tampoco las plataformas que los operan. Es el simple traslado desde un soporte hacia otro considerado mejor, más moderno o simplemente de mayor capacidad: de un disquete a un cartucho, de un cd-rom a un dvd, etcétera. Su objetivo primordial sería resolver la *permanencia* al evitar que los soportes de los documentos lleguen a deteriorarse físicamente, como es el caso del óxido férrico en disquetes o cintas magnéticas. También se propondría actualizar la tecnología de los soportes con miras a poder continuar leyéndolos con el advenimiento de tecnologías más nuevas y, sobre todo, que sigan

estando disponibles, como sería el caso de actualizar archivos contenidos en disquetes hacia CD o de que no existieran los lectores de esos soportes.

Si bien no es ésta la solución total al problema, esta técnica ha sido desde hace décadas la manera más simple de lidiar contra la reducida duración de los soportes y la obsolescencia de la tecnología que los accede subrayo, la tecnología que los accede y “lee”, y no la que los opera intrínsecamente—. En esta técnica, se procede a hacer el recopiado de los registros más o menos de cada cinco a siete años. Algunos podrían pensar ¿pero ello no implica enormes costos? La respuesta es: no del todo. Costos sí; enormes, no. El mayor costo es en tiempo y el de las personas que se ocupan del proceso. Los equipos y soportes prácticamente no son considerables.

Reflexionemos en los siguientes datos: el almacenamiento opto-magnético ha reducido su relación costo/beneficio de una manera impresionante; como ejemplo se encuentra el primer disco duro que existió; la unidad de disco “disk drive” IBM 350, que salió a la venta en septiembre de 1956 con una capacidad de 4.4 Megabytes. En realidad era un “paquete” o conjunto de 50 discos de 24 pulgadas —100 superficies—. Para principios de la década de los ochentas apareció el primer disco duro para PC, el ST-412, más conocido como “Winchester”, el cual tenía una capacidad de diez Megabytes y costaba alrededor de 700 dólares. Diez años después se compraban por la misma cantidad de dinero cien Megabytes; para el año 2000 se podía comprar un Gigabyte —mil Megabytes— y en este momento se compra por la misma suma un Terabyte —mil Gigabytes o un millón de Megabytes—. Esto significa cien mil veces más en veinticinco años por el mismo dinero. Dicho de otra forma, cada ocho años en promedio podemos guardar mil veces más información que por la misma inversión de hace ocho años. Esto significa que dentro de unos siete u ocho años más, compraremos un Petabyte —mil Terabytes o un millón de Gigabytes o mil millones de Megabytes o 10^{15} (diez a la décimoquinta potencia) bytes o un uno con quince ceros de bytes— por los mismos 700 dólares. Otro ejemplo son los discos portables: a principios de los setentas, se guardaban cien kilobytes —un décimo de Megabyte— en un disquete de tres dólares. Diez años después se almacenaba un Megabyte en un disquete por esa cantidad. En los noventas se almacenaban seiscientos cincuenta Megabytes en un cd-rom de treinta centavos de dólar y a la fecha se almacenan cuatro Gigabytes en un DVD de cuarenta centavos de dólar. Cuarenta mil veces en treinta y cinco años; mismo

precio. Proporciones parecidas existen en las cintas y otros dispositivos. Y la tendencia se mantiene.

- El paso siguiente en los procedimientos contra la obsolescencia tecnológica es el conocido como “migración”. Este método, a diferencia del anterior, no se queda en un simple copiado de medios, sino que va más allá; implica el cambio de elementos de equipo o programas, o cambios generacionales de la infraestructura de cómputo. Aquí sí se cambia la tecnología que los opera intrínsecamente. Por ejemplo, cambios de versiones de los documentos tipo *.doc del sistema operativo *MS-DOS* al sistema operativo *Windows*; documentos *.doc de *Office 95* a *Office XP*. Cambio de documentos del formato de *Word Perfect* al de *Office*, o de éste a *Staroffice*. Bases de datos catalográficas en *Unicorn* migradas a *Aleph 500*. En estos ejemplos, la copia no fue sólo hacia otro medio o soporte; implicó también la transformación de formatos de los documentos y las plataformas que los operaban.

El propósito primordial de la migración es el de preservar la integridad de los objetos digitales manteniendo la capacidad de los usuarios de acceder a ellos a lo largo de varias generaciones tecnológicas. La migración incluye al proceso de “recopiado” pero difiere de éste en que no se queda en un simple cambio de medio; la mayoría de las ocasiones implica además cambio de formatos, versiones, sistemas operativos, estándares, etcétera. La imagen copiada difiere del original en lo tocante a su estructura interna, pero no debe diferir en cuanto a su usabilidad. Este proceso por lo general consume mucho más tiempo y recursos que la recopia.

La migración exitosa es definida en mayor medida por el manejo adecuado de lo que se conoce como la “paradoja de la preservación digital”: En general tenemos la sensación de que preservar significa mantener a las cosas sin cambio. Pero en las situaciones donde el documento conlleva cierto grado de tecnología —en especial la tecnología digital—, como ya hemos discutido, si guardamos los documentos sin modificaciones, acceder a esta información se volverá cada vez más y más difícil si no es que imposible. Ambas situaciones combinadas crean esta paradoja. Por un lado está la intención de preservar el documento tan intacto como sea posible, y por el otro, la intención de mantenerlo accesible permanentemente con las mejores herramientas disponibles en ese momento. Se trata de un conflicto entre el “contexto de creación” y el “contexto de uso”, algo que debemos resolver balanceadamente. En general resolver este problema implica contemplar aspectos tales como: contenido del documento, formatos, estilos

y apariencia; contexto del documento; sistema o contexto tecnológico; metadatos y posiblemente el soporte tecnológico de todo esto. Existe una delicada interdependencia entre estos factores que debe ser correctamente interpretada y balanceada para mantener en equilibrios los dos contextos mencionados. Es de vital importancia que, a pesar de los cambios introducidos por el proceso de migración, el documento pueda ser considerado todavía como “íntegro”; esto es, que sigue siendo auténtico, no sólo en su aspecto o soporte, sino en su esencia, en su contenido.

Uno de los secretos de la migración “exitosa” en lo relativo a formatos consiste en apegarse en lo posible a formatos y sistemas no propietarios sino lo más estándar posible: formatos documentales como txt, html, xml, jpg, cdda, etcétera. En este aspecto —más que nunca— el apegarse a estándares y normas a la larga remunera bien.

- La última de las técnicas es la “emulación”; en ella, se pretende replicar la funcionalidad de un sistema obsoleto que ya no tenemos o que ya no funciona. Esto puede entenderse mejor si imaginamos los antiguos juegos electrónicos de video, como los originales de *Atari* o *Nintendo*. Estos pueden ser emulados en una computadora actual tipo PC. No es exactamente el mismo programa antiguo el que vemos en el actual computador. Es un emulador que replica el funcionamiento del anterior para que funcione y se perciba igual. Como segundo ejemplo para entender este concepto, pensemos en el ícono de *MS-DOS* que se observa en los sistemas operativos *Windows*. Cuando lo utilizamos y vemos trabajar en la pantalla ese sistema operativo anterior, en realidad no existe como tal en la computadora; *Windows* se encarga de emular o replicar su funcionamiento para que su uso y percepción por parte del usuario sean semejantes a aquél.

Como puede observarse, el problema real no es la poca duración de los soportes digitales. Como mencioné antes, no están hechos para durar. No quiero decir que su duración no consista un problema; por supuesto que lo es. El punto es que no estriba en ello el problema principal, ya que esa situación puede solucionarse con cierta facilidad con alguna de las técnicas señaladas en este apartado. El verdadero problema con este factor subyace en la *obsolescencia tecnológica* como el mayor obstáculo para el acceso a la información digital en un futuro. Pero todavía mayores enemigos son la desidia, la ignorancia y la falta de atención. El mayor pecado en este aspecto siempre será el de la procrastinación. Benjamín Franklin afirmaba: *Tú puedes demostrar; el tiempo no lo hará*. Debemos estar atentos a los signos del deterioro y a

sus soluciones; si aplicamos correcta y oportunamente las medidas recomendadas funcionarán adecuadamente y el problema se solucionará, aunque sea por unos pocos años. Si actuamos correctamente, *siempre* se puede comprar más tiempo. Pero si dejamos pasar demasiado tiempo muy probablemente tendremos una tragedia documental; en ese caso, no debemos tratar de justificarla después culpando al mensajero tecnológico.

4. FACTORES LEGALES DE LA PRESERVACIÓN

Después de haber observado los factores técnicos de la preservación documental digital es necesario asomarse a los aspectos legales de la misma. Algunas personas no le conceden a este punto la importancia debida y sin embargo es uno de los factores que hoy en día inciden más notoriamente en los proyectos de preservación documental.

El problema nace de la facilidad con que en los tiempos actuales se puede reproducir un documento. En esta era de computadoras, escáneres, impresoras láser, redes y libros virtuales nada más fácil que sacar una copia de un documento digital. Una vez que un documento de este tipo ha sido creado electrónicamente, ¿cómo evitar que rápidamente cientos, miles de copias no autorizadas circulen por la red? En general, a nivel mundial, el libro impreso en papel ha llegado a tal punto de equilibrio que, —aparte las consideraciones de propiedad—, salvo libros muy caros o raros, al público no le es práctico ni rentable, y por tanto no le resulta atractivo, fotocopiar íntegramente un libro en lugar de adquirir el original, y mucho menos atractivo resultaría hacer una copia mecanografiada del mismo.

El caso de un libro en forma electrónica es distinto; las maneras y soportes para copiarlo son variadas, y los costos de ello son prácticamente nulos. Cuando el propósito de un autor o editor es maximizar la distribución, —con pocas o nulas miras económicas—, la tecnología digital les permite hacerlo de manera muy eficaz y rentable. Pero cuando la edición tiene como propósito la venta y con ello una utilidad económica, el panorama cambia. Muchos editores se han detenido por mucho tiempo y se manejan, —con toda razón— con suma cautela ante este hecho para lanzar documentos digitales al mercado.

Lo que ha sucedido en las últimas décadas debido a lo anterior es que los editores, —quienes no han encontrado mecanismos de comercialización realmente innovadores para sus productos, amenazados por el creciente número de editores no autorizados, así como la tradicional tendencia de los editores de arremeter de tiempo en tiempo contra los derechos de libre acceso

de los usuarios y estirar a su favor las legislaciones para obtener otros ingresos extra—, han incrementado cada vez más y más las cortapisas legales a las posibilidades de hacer una copia documental por parte del público bien intencionado y de múltiples instituciones culturales, entre ellas bibliotecas, archivos, museos, etcétera.

En efecto, tal como en tiempos antiguos, muchos editores piensan que la mejor forma de proteger las obras contra daño o sustracción es imponer terribles amenazas y castigos para quienes lo hicieren: desde maldiciones, excomuniones y anatemas en la antigüedad hasta leyes y reglamentaciones, multas y cárcel, —entre otras medidas—, han sido lanzados contra los infractores por más de cuatro milenios. O peor aún, simplemente la solución ha sido y todavía es, —por desgracia tiende a repetirse con chocante frecuencia—, el evitar prestar las obras. Por un milenio y medio ha probado también ser medida inútil para resolver el problema, lapso que incluye los tiempos actuales.

Los editores, además de no resolver su problema de ventas, han creado otro grave problema de fondo: la falta de preservación y pérdida de materiales documentales. Para ilustrar esto tomemos dos recientes estudios encargados por la Biblioteca del Congreso de los E.U.A. al CLIR —Consejo de Recursos en Bibliotecas e Información— o *Council on Library and Information Resources*. En ellos se reporta este fenómeno. En particular, se resalta el hecho de que en 1972, el Congreso de ese país extendió una protección federal especial para los registros fonográficos hechos en los E.U.A. a partir de ese momento; pero también incluyó una serie de parches y enmiendas legales a lo grabado antes de ese año. Gracias a ello, —o mejor dicho, desgraciadamente por ello— nada pasará al dominio público sino hasta el año 2067. De acuerdo con esta legislación, un registro de audio puede ser copiado si el original se encuentra en un formato “obsoleto”, pero de acuerdo con el estudio, obsoleto significa en esta legislación que el dispositivo para reproducirlo no se produce más comercialmente. Por desgracia, las tornamesas para discos de 33, 45 y 78 r.p.m. se producen todavía y esto hace que no se permita hacer copias de todos estos materiales para preservación; por eso gran parte de esos materiales no han sido reeditados en soportes más modernos [Besek, 2005].

Por esta razón, el Congreso de los E.U.A. emitió el *Acta Nacional de Preservación de Grabaciones del Año 2000* en la cual ordenó un estudio del derecho de copia y otras normatividades aplicables a las grabaciones de audio. Como resultado de esos estudios, se concluyó que: .

..con excepción de unos cuantos registros de algunas compañías cuyos acervos han sido abandonados o pasados al público, no existe el dominio público de grabaciones de audio en los E.U.A.ya que sólo el propietario del derecho de copia

puede legalmente hacer que los antiguos registros se vuelvan disponibles, los registros históricos están en grave riesgo tanto de pérdida física como de pasar desapercibidos de la memoria de audio del país [Brooks, 2005, p. 6].

Otro de los estudios concluye:

existe una clara evidencia de la necesidad de actualizar las leyes de propiedad para obtener ventajas de la tecnología digital para preservar y hacer accesible la mayor parte del acervo nacional de registros de audio [Besek, 2005, p. vii].

Los estudios también abarcan postulados iguales para los registros filmicos y de video, donde el tétrico panorama no difiere mucho del presentado con respecto a los registros de audio.

Este ejemplo sirve para ilustrar un hecho que empieza a tomar tintes de problema grave: los ajustes hechos a las legislaciones a nivel mundial, promovidas por el aumento de piratería de editores ilegales, han desembocado en una serie de pegotes, parches y enmiendas legales hechas a tontas y a locas, los cuales responden a impulsos del momento o a presiones e intereses de grupos, sin un entendimiento ni una visión integral y a largo plazo de lo que el problema representa, y sin un análisis de las repercusiones que ello implicará y de que los males que se crean son inmensamente mayores que los beneficios. Las consecuencias pueden notarse ya:

- El problema a nivel mundial de la piratería por parte de editores ilegales y sin escrúpulos no se ha resuelto; ni siquiera tiende a disminuir. Por el contrario, tiende a incrementarse, pese al aumento de legislaciones y el endurecimiento de sanciones al respecto.
- Las restricciones para el público que *no* se dedica a la piratería y que obtiene copias ocasionales —y las más de las veces legales— empiezan a ser molestas para ese sector y ocasionan cada vez más confrontaciones entre ambos grupos de intereses, así como un incremento en el menosprecio que el público siente por los derechos del editor.
- Se ha creado y se incrementa un enorme vacío de responsabilidad en la preservación documental, lo que pone ya enormes cantidades de material en riesgo de perecer y con lo cual incurrimos en una grave responsabilidad generacional.

De estos tres fenómenos, obviamente el que tiene que ver más con la preservación es el tercero, y por tanto abundaremos en esta idea.

4.1 El Depósito Legal

Para muchos autores y editores, el depósito legal de materiales digitales es considerado como un riesgo potencial para sus derechos de propiedad. Éste es un fenómeno nuevo; con los materiales en soportes “tradicionales” nunca fue considerado así. Buena parte de este problema nace de las ambigüedades que se han introducido alrededor de las leyes de propiedad intelectual y las del depósito legal; los parches y pegotes ya comentados. En buena medida los ajustes legales fueron simples extensiones de lo establecido para materiales en soportes tradicionales, sin tomar en cuenta que en realidad los materiales documentales digitales *requerían de nuevos modelos de protección* que dejasen perfectamente claras cuáles eran las semejanzas y cuáles las diferencias con los anteriores, y que debe buscarse siempre el equilibrio entre las dos protecciones mencionadas: la de los derechos del autor y la de preservación de los materiales y el derecho a su acceso. Si estas semejanzas y diferencias no se entienden será muy difícil lograr acuerdos entre los diversos actores, y mucho menos legislar acertadamente acerca de ellas. Por tanto, continúo mi análisis con el depósito legal; en especial el *depósito legal digital*.

Como puede intuirse, la idea de tener un ejemplar de una obra en lugar seguro para la posteridad no es nueva. Como muchas otras ideas alrededor de los libros y los documentos data de hace siglos, pero no en todas las épocas y lugares se formalizó; por lo general se trataba de un acto aislado de magnanimidad o visión de un monarca. La importancia histórica del depósito legal radica en que gracias a este acto, las obras publicadas en una época determinada eran, y son, preservadas por largo plazo en un espacio nacional reconocido y comprometido con ello. La legislación y el carácter obligatorio del depósito legal aseguran la recolección y la preservación de la herencia nacional de publicaciones, y por ende, de una importante parte de la riqueza cultural de un país.

Las disposiciones de depósito legal y las de *copyright* se entrelazaron gradualmente durante los siglos XVIII y XIX. En esa época en la mayoría de los países era necesario hacer el depósito para obtener la protección del derecho de propiedad; de otra manera los editores no quedaban protegidos. Esta situación terminó con el *Convenio de Berna* de 1886 acerca del *copyright*, en el cual se acordó que la protección de ese derecho no podía quedar limitada por ninguna formalidad, y entre ellas se incluyó al depósito legal. Este principio permeó gradualmente a la mayoría de las legislaciones acerca del depósito legal en todo el mundo durante las primeras décadas del siglo XX; por ello, las disposiciones acerca del depósito continuaron en casi todos los países, pero cada vez más de forma disociada del derecho de propiedad, como fue el caso de México.

Con el correr de los siglos, como era de esperarse, los requisitos y características del depósito han evolucionado. El concepto se ha expandido a otro tipo de publicaciones además de los libros; por ejemplo, los materiales audiovisuales. En Viena se creó el primer archivo audiovisual en 1899, el cual partió expresamente de ese tipo de fuente documental. Ha variado también el número de copias que deben depositarse, así como la institución o instituciones responsables de la recepción del depósito y del acceso al material. Y también surgió la elaboración de bibliografías nacionales como subproducto del depósito, etcétera. En cada nueva época, las características de los tiempos han impuesto más y más presiones y cambios sobre la reglamentación del depósito legal. Con las nuevas características introducidas por las publicaciones electrónicas esta normatividad enfrenta nuevamente formidables retos para poder actualizarse frente a este tipo de publicaciones; retos que no tan sólo provienen del campo técnico u organizacional, sino primordialmente del mismísimo punto de vista legal.

Las publicaciones digitales adquiridas por las bibliotecas a través de suscripción o licenciamiento –por ejemplo las revistas o los libros electrónicos– son un ejemplo muy ilustrativo de esta paradoja que surge ante la necesidad de proteger o preservar los materiales a la vez que la necesidad de proteger los derechos de propiedad. En esta estructura actual, en general, los documentos adquiridos en esta forma son consultados vía la *Web* en un servidor propiedad del editor o de un concesionario de él. La biblioteca adquiriente, por lo general, no se queda con copias de los materiales documentales. Bajo este esquema, la institución que posee los materiales no tiene la responsabilidad de preservarlos a largo plazo y puede darlos de baja cuando quiera; en cambio, la institución encargada de preservarlos no posee esos materiales. Esta paradoja debe ser resuelta de alguna forma con algún acuerdo concertado entre autores, editores, bibliotecarios, legisladores, etcétera.

A raíz de la aparición de este tipo de publicaciones, licenciamientos, etcétera, han empezado a surgir las siguientes preguntas, relacionadas con la necesidad de la preservación documental:

- ¿Quién debe preservar estos materiales: el autor, los editores/productores, ciertas bibliotecas designadas o la biblioteca nacional del país?
- ¿Cómo preservarlos por largo plazo cuando muchas veces la institución encargada de preservarlos no posee copias ni derechos sobre los programas o plataformas digitales que pueden explotar los documentos?
- ¿Esas bibliotecas ¿pueden prestarles los documentos al público? Si la respuesta es sí, entonces ¿cómo y en qué condiciones? Si la respuesta es no, entonces ¿para qué preservarlos?

Trataré de dar respuesta a estas preguntas. De acuerdo con el *Taller acerca de temas relacionados con colecciones nacionales de publicaciones electrónicas por depósito*¹⁴ celebrado en 1995, la necesidad de que hubiese depósito legal de las publicaciones electrónicas fue un acuerdo entre bibliotecarios y editores. Hubo consenso que de no ser así, las publicaciones estarían en riesgo en un cierto plazo. En efecto, pretender que los autores o editores deberían tener la responsabilidad de preservar los documentos por largo plazo es absolutamente impensable. Por su misma naturaleza estas personas no pueden garantizar la futura disponibilidad y continuidad del acceso.

Si bien hoy en día muchos editores desean poseer la información y permitir el acceso a sus publicaciones en sus propios sitios *Web*, está comprobado que esto es impulsado por el deseo de tener el control total -legítimo por cierto- de sus publicaciones. Pero después de cierto lapso este deseo se reduce y acaba por ser nulo al cabo de un tiempo, pues ello no les es ni práctico ni económicamente rentable. Un ejemplo claro lo tenemos ahora con las publicaciones periódicas técnicas y científicas, las cuales en su gran mayoría son ofrecidas directamente por las casas editoras en sus propios sitios *Web*, con su propio computador y empleados propios. Si observamos un poco más de cerca veremos que la mayoría de estas casas no garantiza que seguirán ofreciendo estas revistas más allá de unos cinco años después de su publicación o, en el mejor de los casos, diez. Es claro que para ellos, después de esos plazos, que ya tienen muy bien estudiados, las ganancias de esos años retrospectivos de las revistas son mínimas mientras que los costos se mantienen. Pero entonces ¿quién las va a preservar y en su caso a distribuir? Como mencioné, es claro que pasado cierto lapso se necesitará una organización dónde depositarlas y que el depósito en bibliotecas u organizaciones semejantes debería ser indispensable y algo que debe hacerse de común acuerdo entre editores, bibliotecarios y usuarios.

4.2 *El derecho de Copia o Copyright*

Después del análisis del depósito legal, mecanismo por excelencia de preservación documental para bibliotecas nacionales y otros repositorios de carácter nacional, es necesario hacer ahora un análisis de los mecanismos que tienen que ver con otro tipo de bibliotecas y archivos: escolares, universitarios, públicos, etcétera, así como de empresas y público en general. En esencia, el mecanismo son las leyes acerca del derecho de autor y de propiedad o *copyright*. Un análisis detallado al respecto se hizo en el libro *Bibliotecas y*

14 *Workshop on Issues in the Field of National Deposit Collections of Electronic Publication*, 1995, European Commission, DG Information Society, Luxembourg, December 18, 1995, disponible agosto 2008 en: <http://cordis.europa.eu/libraries/en/depo-rpt.html>.

Publicaciones Digitales a lo largo del capítulo 5 “Publicaciones Electrónicas y Derechos de Autor y de Propiedad”,¹⁵ por lo que únicamente haré un pequeño extracto aquí para poner el tema en contexto y poder continuar el análisis.

Es importante además resaltar que las legislaciones de derechos autorales tienen una jurisdicción, obviamente dentro del país que las emite, –en nuestro caso México–, y en algunos casos acuerdos internacionales, que no universales. Dada la naturaleza sin fronteras de la *Internet*, el caso se complica ya que con frecuencia es muy difícil saber cuál es la ley, jurisdicción o acuerdo que debe aplicarse. Por lo mismo, analizaremos las características más relevantes de ambos tipos de legislaciones para poder entenderlas en contexto. Por la naturaleza de su trabajo, los bibliotecarios están más relacionados con el *derecho de copia* o *derechos patrimoniales* que, como ya he mencionado, tienen que ver con la explotación y copia de una obra dada. Si profundizamos un poco más, ya dentro del ambiente de las bibliotecas digitales es necesario que el bibliotecario profesional moderno tenga muy claro qué puede copiar del material digital y en cuáles circunstancias. De inicio, debe saberse que de acuerdo con la ley mexicana vigente, –y con muchas otras legislaciones de países de origen legislativo semejante–, las publicaciones que se pueden copiar sin restricción de derechos patrimoniales o de copia son:

- Las que son del dominio público, ya sea por su propia naturaleza, porque sus derechos han prescrito o porque así lo ha decidido expresamente el autor. En este sentido el bibliotecario profesional debe estudiar con cuidado los análisis y tablas establecidos al efecto, ya que las prescripciones cambian de acuerdo con leyes internacionales o de cada país y se mueven con el tiempo en función de extensiones y prórrogas establecidas sobre las obras.
- Las siete excepciones permitidas en la ley mexicana en su capítulo II, artículo 148. Véase apartado 5.1.1 de la obra *Bibliotecas y Publicaciones Digitales* o en la página de legislaciones mexicanas del Instituto de Investigaciones Jurídicas de la UNAM.¹⁶

De acuerdo con la ley norteamericana del *copyright*, las condiciones en que se puede copiar legalmente un documento, además de aquellos que son del dominio público, están establecidas de acuerdo con lo siguiente:

15 Juan Voutsás M., 2006, *Bibliotecas y Publicaciones Digitales*, Centro Universitario de Investigaciones Bibliotecológicas, UNAM, 343 p. Capítulo 5, “Publicaciones Electrónicas y Derechos de Autor y de Propiedad”, pp. 193-234.

16 Instituto de Investigaciones Jurídicas, Universidad Nacional Autónoma de México (UNAM), 2000, *Ley Federal del Derecho de Autor*, disponible junio 3, 2008 en: <http://info4.juridicas.unam.mx/ijure/fed/139/default.htm?s=>

107. Limitaciones a los derechos exclusivos: uso ético

No obstante las disposiciones de las secciones 106 y 106A, el uso ético de un trabajo protegido con copyright, las que incluyen aquellas como la obtención de copias o registros sonoros o por cualquier otro medio especificado en estas secciones, aquellos con propósitos tales como la crítica, comentario, reportaje, enseñanza –incluidas las copias múltiples para uso en clase–, estudio escolar o investigación no constituyen una infracción del derecho de copia. Para determinar si el uso hecho de una copia de un trabajo cae dentro del “uso ético” deberán considerarse los siguientes factores:

- El propósito y carácter del uso, considerando si éstos son de naturaleza comercial o con fines educativos no lucrativos.
- La naturaleza del trabajo protegido. –¿está publicado o no? ¿es obra factual o artística?–
- La cantidad y relevancia de la porción copiada, en relación a la obra en su conjunto.
- El efecto de ese uso sobre el mercado o valor potencial de la obra copiada.¹⁷

Además de lo establecido en la ley, pueden utilizarse además las reglas de “uso ético de la información” –*Fair Use*–, y para los materiales digitales también pueden utilizarse ahora los materiales distribuidos bajo el principio legal de “Creative Commons”¹⁸, el cual se basa en el principio de “algunos derechos reservados” y ha sido diseñado expresamente para documentos digitales. Ambos principios han sido tratados en el capítulo 5 de la obra “Bibliotecas y Publicaciones Digitales” por lo que ya no abundaré más en ellos.

El advenimiento de las publicaciones electrónicas ha transformado notoriamente el contexto de los derechos de copia inherentes a las publicaciones y es un grave error pretender simplemente extrapolar lo anteriormente establecido para las publicaciones en soportes “tradicionales”, –por ejemplo en papel– ya que ello induciría a caer en lagunas o errores en el manejo adecuado de los derechos de copia en una biblioteca que contenga o acceda a este tipo de publicaciones. Y esta advertencia no es tan sólo para los bibliotecarios: los legisladores tienden mucho a lo largo del orbe –México no es la excepción– a modificar las leyes de protección autoral y tratan a las publicaciones digitales simplemente como si fueran una extensión de los soportes

17 “Copyright Law of the United States of America and Related Laws Contained in Title 17 of the *United States Code*”, disponible junio 3, 2008, en: <http://www.copyright.gov/title17/92chap1.html>

18 Sitio Oficial del Proyecto “Creative Commons”, disponible junio 3, 2008 en: <http://www.creativecommons.org>

tradicionales. Grave error; los resultados son parches, pegotes, masas amorfas que crean muchos más problemas de los que pretenden resolver, y en realidad agravan el problema de los derechos de propiedad. Pero peor aún, están creando un enorme riesgo y mayores problema para la preservación de documentos digitales.

Esta falta de precisión en las legislaciones autorales, con respecto a nuevas variantes que contemplen de manera adecuada las características de las publicaciones digitales, introduce serias lagunas y riesgos enormes para la existencia misma, a largo plazo, de las colecciones y sus usuarios, y por lo mismo deben ser reflexionadas y enmendadas esta vez íntegra y correctamente por los responsables de esas legislaciones con ayuda de los responsables de las colecciones documentales. Es de capital importancia buscar modelos que les restablezcan el equilibrio de la protección a los autores junto con el derecho de los usuarios a la información. Es absolutamente vital rescatar los principios de jurisprudencia de las legislaciones mundiales en el sentido de que: “el propósito del ‘copyright’ es promover el progreso del conocimiento y darle al autor de una obra un incentivo económico para crear nuevas obras”.

No obstante lo anterior y agravando el asunto, algunas legislaciones en diversas naciones han extendido más y más las protecciones de sus leyes de derechos de autor mucho más allá de lo usual en materiales impresos, y con ello desbalanceado totalmente el equilibrio que debe existir entre los derechos de los autores y los de los usuarios. Esta preocupación de que los derechos de los usuarios no pueden estar en desbalance con los de los autores, ha sido recogida claramente por las bibliotecas y sus organizaciones. La Federación Internacional de Asociaciones e Instituciones Bibliotecarias –IFLA– establece claramente estas posturas en su documento “Position Paper on Copyright in the Electronic Environment” sancionado por sus comités en 1996 en la 62ª Conferencia Mundial de esta organización. [International Federation of Library Associations and Institutions, 1996]. Esta postura ha sido reforzada en el Manifiesto de la IFLA/UNESCO *Sobre Internet : Directrices*, en particular en el Tema 7 de las Directrices “Barreras”, inciso “c”, “Propiedad Intelectual”:

- Los recursos de información creados con financiación pública deben estar en el dominio público y permanecer en él.
- Ante legislaciones restrictivas sobre propiedad intelectual en el ámbito digital y en relación con la gestión de derechos digitales, las bibliotecas deben promover alternativas legítimas a las formas actualmente existentes de los derechos de autor y de copia, como las licencias del tipo Creative Commons que aumentan, en vez de restringir, el acceso a la información.
- Los bibliotecarios tienen la responsabilidad de abogar por condiciones de

propiedad intelectual y derechos de reproducción que faciliten la preservación de los materiales digitales; y habría que animar a los propietarios de los derechos a que reconozcan su obligación de asegurar la disponibilidad a largo plazo de los recursos en línea¹⁹. [IFLA / UNESCO, 2006].

¿Cómo afectan los derechos de propiedad autoral a la preservación? Las legislaciones autorales, independientemente de su origen, tipo o país, incluyeron durante los pasados doscientos años, –*la era de los soportes* “tradicionales”– los elementos suficientes que les permitían a las bibliotecas y archivos cumplir sus funciones, entre ellas las muy relevantes de distribuir la información y la de preservarla. Entre esos elementos distinguimos:

- La “Doctrina de la Primera Venta”.
- La posesión física de los documentos. Es decir la biblioteca poseía físicamente una copia del libro revista, periódico, etcétera.

La “Doctrina de la Primera Venta” es de singular importancia, aunque debido a su redacción legal pasa casi desapercibida. Esa doctrina se creó hace casi dos siglos y fue rápidamente adoptada desde entonces y hasta la fecha por casi todas las legislaciones autorales del mundo. Establece que un autor puede oponerse a que su obra se venda o distribuya por primera vez, pero una vez que él ha consentido a ello el derecho de oposición cesa, y el autor no puede impedir que se distribuyan ejemplares de la obra. Establece además que esa copia que adquirimos es de nuestra propiedad y gracias a ello todos podemos hacer con ella lo que deseemos: regalar, prestar o vender un libro, una revista, un disco u otra publicación semejante que hemos adquirido y deseamos transferirle a otra persona o institución, sin caer en faltas a la ley. Podemos además tirarla, subastarla, etcétera, No obstante, queda entendido que al venderla o regalarla entregamos la copia por la cual ya habíamos pagado y que al transferirla a otra persona ya no tenemos esa copia en nuestro poder, y no podemos hacer más copias ni otros usos de ella. Ésa es la única limitante dentro de la doctrina: no copiarla. De hecho, ha sido por siglos la esencia misma de cómo compartimos información en este mundo.

Por siglos hemos desarrollado y arraigado la idea de que si nos gusta algo que compramos para leer, podemos recomendarlo, prestarlo o regalárselo a un amigo o colega,

19 IFLA, International Federation of Library Associations, 2006, *Manifiesto de la IFLA/UNESCO Sobre Internet : Directrices*, Compilado y Editado por Seidelin, Susanne, Directora de la Oficina IFLA/FAIFE, septiembre 2006, Página Oficial de la IFLA, disponible agosto 2008 en: <http://www.ifla.org/faife/policy/iflastat/Internet-ManifiestoGuidelines-es.pdf>

y además si decidimos que ya no deseamos esa obra, podemos tirarla o venderla. Estos principios son tan importantes en nuestro quehacer intelectual que debemos tenerlos presentes siempre y encontrar la manera de trasladarlos al entorno de las publicaciones digitales. [Soules, 2002].

Cuántas veces sucede ya en la actualidad, que los usuarios no pueden conservar para sí una copia de un artículo o documento necesarios para una tarea, trabajo o investigación, incluso después de haber pagado por ella.

Esta doctrina es la que les ha permitido también a las bibliotecas prestar libros desde que hay leyes autorales. Si las bibliotecas tuvieran que pagar una regalía cada vez que le prestan un libro a un usuario seguramente tendrían que cobrársela a éste, y las bibliotecas y la educación no hubieran sido lo que son hoy. Es cierto que en algunas épocas y países, sí llegó a existir una “regalía” por el préstamo de un libro. Afortunadamente, se establecieron también muchas bibliotecas en países nacientes en el siglo XVIII y XIX con el concepto de “gratuidad”, concepto que permeó a prácticamente todas las bibliotecas a nivel mundial. Nótese bien que si hoy tuviéramos el modelo del pago por el préstamo, la única y mínima diferencia entre una librería y una biblioteca sería que una vende y la otra renta los libros, y las bibliotecas tal como las conocemos *no existirían ya*. Esta característica no ha sido gratuita para las bibliotecas; desde hace varias décadas las suscripciones a revistas son visiblemente más caras a bibliotecas que a personas, y frecuentemente sucede esto también con otras publicaciones.

Nótese que muchas bibliotecas pagan derechos extras adicionales por reproducir para el público una obra o parte de una obra, pero no por prestársela. Por la obra original la biblioteca ya pagó: ésa es la esencia misma del préstamo bibliotecario que no ha quedado adecuadamente plasmada en las legislaciones autorales “digitales” y que tantos problemas causa todavía hoy en día.

El segundo punto enunciado –la posesión física de los documentos– es clave en la preservación documental. ¿cómo preservar algo que no se posee físicamente? El punto toral del problema en la actualidad es que hemos migrado de una estructura basada en las leyes autorales. (en donde ya hemos observado que las funciones de la biblioteca –distribución y preservación–, estuvieron garantizadas y en armonía con los editores y los usuarios) hacia una estructura basada en contratos de licenciamiento de acceso, en donde se perdió toda esa armonía que existía entre legislación-bibliotecas-editores-usuarios. En estas nuevas estructuras de acceso a la información digital la biblioteca pierde la posesión de los materiales porque ésta reside en poder de los editores o distribuidores, y prácticamente no existe doctrina de primera

venta. El hecho de que las bibliotecas no posean físicamente los materiales —ya que obtienen por su dinero un contrato que sólo otorga una licencia de acceso— conlleva, a la larga, a que la biblioteca no podrá conservar, preservar ni restaurar esos materiales.

Si existiese la obligación moral y legal hacia los editores o distribuidores de preservar su material documental para el futuro y darle así acceso al público, no habría problema. Pero como ya se ha comentado, no ha estado ni está en los objetivos de la inmensa mayoría de las empresas editoriales o distribuidoras el propósito de gastar recursos en preservar ese material para el futuro. Y en realidad no tendría que estarlo: no es su función ni su responsabilidad social.

El fin de la posesión física de los materiales documentales por parte de las bibliotecas, su nula preservación a largo plazo por parte de los editores y el fin de la doctrina de la primera venta le crean problemas graves al futuro de la preservación documental a nivel mundial, y por ello debemos seguir buscando equivalentes funcionales y legales para esa preservación. Estos puntos permanecen como sumamente importantes y delicados en cuanto a las colecciones digitales. Los usuarios han creado derechos y costumbres de cómo usar una obra y a obtener sus “copias incidentales temporales”, —como ha sido ratificado en algunas legislaciones actuales— a lo largo de muchísimos años, y es imposible ahora tratar de imponer que cuando se paga por una obra sólo se adquiere el derecho de ver u oír esa obra por un lapso finito de tiempo, o en un sólo lugar, y que es exclusivamente para sus ojos y no puede ser compartida o poseída en modo alguno. Eso simplemente va en contra de la naturaleza hombre-libro y algunos de sus derivados. Ningún modelo comercial, tecnológico o legal de alta restricción, entre los establecidos a la fecha, ha funcionado ni parece prometer algo real en el futuro; ello se debe a que esos nuevos modelos carecen del “equivalente funcional digital” a estos principios y a la doctrina de la primera venta en los documentos electrónicos. Esto, que aparenta ser algo trivial y sin importancia, implica una serie de graves desquiciamientos en la mismísima esencia de la transferencia de información, sobre todo a largo plazo. Por esta razón los riesgos de pérdida masiva de materiales y colecciones documentales crecen día a día al crearse estos materiales documentales y siguen sin existir los mecanismos legales que permitan su preservación. Pero si no existen los principios que rijan su preservación, ni las adecuadas y pertinentes definiciones acerca de quién debe preservarlo ¿cómo se va a poder preservar este patrimonio documental? ¿quién lo va a hacer?

En realidad enfrentamos un problema de conciencia y sensibilización que en su mayor parte es de índole legal, aunque por supuesto incluye sus problemas técnicos y sociales. A nosotros, los profesionales de la bibliotecología,

nos corresponde esta primera parte: la de concientizar y sensibilizar: primero a nosotros mismos; y después a los tomadores de decisiones y planeadores, y también a los editores y distribuidores, y finalmente –y no por ello menos importante– a los legisladores, ya que es en ese nivel en el que deben establecerse los elementos de equivalencia funcional que vuelvan a equilibrar los derechos de los autores, editores, bibliotecas y usuarios en forma armónica, y que permitan crear responsables y los mecanismos de esa preservación digital a largo plazo; sin pegotes ni desequilibrios, y sin sensaciones de abuso de un sector por parte de otro.

5. FACTORES DOCUMENTALES DE LA PRESERVACIÓN

La definición más sencilla de biblioteca que he visto alguna vez –hace ya tanto tiempo que no puedo recordar dónde–, y que no obstante su sencillez es absolutamente válida afirma que

los tres elementos mínimos para que exista una biblioteca son: una colección documental, con un sistema de orden y registro, y un conjunto de usuarios interesado en esa colección.

Ésta es una definición a lo Joan Miró sin duda: tres pinceladas para pintar de cuerpo entero una biblioteca, desde la más pequeña hasta la más grande. Obviamente si deseáramos especificar el tipo de biblioteca, su alcance, etcétera, tendríamos que agregarle elementos a esta definición; esto es incuestionable. Pero la esencia misma de las bibliotecas, su núcleo irreductible reside en esta sencilla definición. Es válida incluso para las actuales bibliotecas digitales, y sin duda marca claramente la diferencia entre la información amorfa existente en la *Internet* y las verdaderas colecciones que existen dentro de ella.

Subrayo uno de sus elementos: “un sistema de orden y registro”. En efecto, no puede concebirse una biblioteca sin los mecanismos que permitan registrar y ordenar la información –obviamente para poder recuperarla después– y es uno de los elementos irreductibles; una biblioteca no puede existir sin ellos.

Una de las grandes ventajas que nos ha traído la red en su conjunto es que *contiene una enorme cantidad de documentos*; es una vasta colección documental sin duda. Uno de los grandes problemas que también nos ha traído la red en su conjunto es precisamente que *contiene una enorme cantidad de documentos*. Su gran volumen y su rápido crecimiento son a la vez una ventaja y un problema.

Y en efecto, hay o puede haber un conjunto de usuarios interesados en esa información; es un universo en expansión. Pero en una inmensa proporción esa información, ese universo, no tiene sistema de registro y orden. O éste es tan primitivo que no permite la recuperación coherente del material. Simplemente se recuperan miles y millones de palabras y registros por medio de la “fuerza bruta” de las computadoras. ¡Cuántas veces hemos hecho alguna búsqueda en la red para encontrarnos con millones de supuestas referencias, pero en realidad poco o nada relevantes acerca del material que buscamos! Ello se debe precisamente a los deficientes o nulos mecanismos de registro y orden que todavía se usan en la mayoría de los reservorios documentales que existen dentro de la red; de ahí la inferencia a la existencia de ese universo entrópico de la red. Ya hemos mencionado, entre otros autores, a Clifford Lynch y Michel Gorman [Lynch, 1997] y [Gorman, 1998], reconocidos académicos estudiosos de la información documental en la red, cuyos trabajos destacan por sus interesantes disertaciones acerca de las razones por las cuales la red no puede considerarse –vista como un todo– como una enorme biblioteca global. Puede abundarse además en el excelente artículo de Mark Y. Herring, decano de los servicios bibliotecarios en la Biblioteca Dacon de la Universidad de Winthrop, Rock Hill, Carolina de Sur, en la unión americana, denominado “Diez Razones por las que Internet no Sustituye a una Biblioteca”.²⁰

Imaginemos por un momento una gran biblioteca tan grande como podamos, en un inmenso edificio que tiene una innumerable cantidad de salones con cuantiosas estanterías que contienen una magnífica colección: –enorme, interesante, completa–, pero donde los libros estuvieran acomodados por color de las cubiertas, o por tamaño, o por la persona que los compró, o por la fecha en que llegaron a la colección. Tratar de encontrar algo ahí sería imposible y frustrante, a pesar de lo magnífico de la colección. Entre mayor su volumen, mayor el problema. Ése es precisamente el obstáculo con que nos topamos en la actualidad en muchas colecciones digitales: gran volumen pero una falta de registro y orden, o un orden tan primitivo, heterogéneo y elemental que por lo mismo resulta inútil al momento de tratar de recuperar algo.

Por esta razón: la de estar conscientes de que una vez que la colección documental llega a un cierto tamaño es indispensable registrarla y ordenarla, si es que alguna vez deseamos poder recuperarla, es por la que desde las

20 Mark Herring Y., 2001, “10 Reasons Why the Internet is no Substitute for a Library”, en: *American Libraries*, April 2001, pp. 76–78, disponible también agosto 2008 en: “American Libraries OnLine : News and Updates from the Magazine of the American Library Association”, Página Oficial de la ALA, <http://www.ala.org/ala/online/selectedarticles/10reasonswhy.cfm>.

primeras bibliotecas de la antigüedad se comenzaron a desarrollar e incluir estas ayudas, y se le agregó a la colección un mecanismo que permitiera darle ese orden que, a su vez, permitiera su posterior localización. Este mecanismo son los metadatos: datos acerca de los datos; registrar con método para poder ordenar y recuperar después. Los encabezados grafiados en los costados o principios de las tabletas de arcilla cuneiformes eran metadatos; las tarjetas en los canastos con papiros en Alejandría eran metadatos; las bibliografías medievales eran metadatos; las fichas catalográficas en cartón y los registros en formato MARC del siglo XX son metadatos. Aunque el concepto es varias veces milenario, fue en el verano de 1969 cuando Jack Myers acuñó el neologismo que usamos actualmente –el metadato– para describir conjuntos de datos acerca de la información. Él lo usó en una línea de productos que él llamó su metamodelo. Posteriormente fundó una empresa con ese nombre: *The Metadata Company*. El término se abrió paso hacia el dominio público e inclusive amplió y enriqueció su idea original hasta la acepción moderna y mucho más compleja que hoy conocemos.

Antiguos o modernos como la idea y el término puedan ser, lo importante es su esencia: El metadato conlleva la esencia de orden dentro del universo de las publicaciones una vez que se hacen numerosas, desde las tabletas de arcilla hasta los esquemas XML y probablemente en un futuro en los esquemas semánticos. La parte realmente valiosa de la información dentro de la *Web* es aquella que ha sido estructurada de manera sistemática y coherente con metadatos adecuados. Es la que permite encontrar y recuperar de forma efectiva la información una vez guardada. Por ello las bibliotecas y archivos digitales, –me refiero a las verdaderas bibliotecas y archivos– que se encuentran dentro de la *Web* son los que de manera efectiva pueden brindarnos información dentro del caótico mundo de la *Web*. Cuando digo las verdaderas bibliotecas y archivos no me refiero tan sólo a las que así se ostentan, sino a todos aquellos acervos que tienen estos tres principios básicos: acervo, buen sistema de orden y usuarios potenciales. El nombre es lo de menos. Lo importante es que realmente cuenten con lo esencial para almacenar y recuperar acervos.

De estas reflexiones surge lo que he denominado la paradoja de las bibliotecas digitales, la cual en realidad sólo es una aparente paradoja: “La *Web* como tal no es una biblioteca digital, pero todas, absolutamente todas las bibliotecas digitales están contenidas dentro de la *Web*”. Y digo aparente porque si se analiza con cuidado podrá deducirse que no es una paradoja como a simple vista podría suponerse: es una cuestión de saber acomodar correctamente las proporciones de lo que las bibliotecas digitales –reitero, las verdaderas bibliotecas– representan dentro de la red. Por supuesto esto es también extensivo a los verdaderos archivos que hay en la red. Y como hemos podido

reflexionar, los metadatos son el elemento que hace esa diferencia y por tanto son de crucial importancia.

Repasemos, los metadatos son datos estructurados que definen a otros datos con objeto de identificarlos, describirlos, clasificarlos y localizarlos; nos brindan información sobre diversos atributos del dato que hay que definir: su contenido, sus aspectos formales o físicos: nombre, tamaño, tipo del dato, fechas, idioma, campos que lo conforman, propiedad. Sus aspectos contextuales: calidad o grado de preservación, autenticidad, cualidades, asociaciones, condiciones de uso, etcétera. En suma, información acerca del dato y sus características, contexto y cualidades. Los metadatos en la actualidad pueden asociarse a cualquier tipo de documento o —más formalmente—, a cualquier objeto documental. El “dato” al que se asocian puede ser cualquier tipo de documento formal o tradicional: libro, revista, manuscrito, artículo, partitura, periódico, tesis, mapa, música, imagen, video o película, etcétera. Es decir, cualquier documento, en su acepción más amplia, y pueden ser usados tanto en soportes *tradicionales* como digitales. Los metadatos no tienen por fuerza que estar asociados a documentos digitales, aunque no hay duda de que es en este ambiente de registros electrónicos donde puede explotarse su potencial en toda su enorme dimensión. Los metadatos de un objeto documental pueden ser simultáneamente el contenido de otro objeto documental. Una bibliografía, por ejemplo, es a la vez obra y metadato de otras obras.

Los metadatos comienzan obviamente con la simple descripción de un objeto documental. Los profesionales de bibliotecas y archivos han utilizado secularmente los principios de catalogación o descripción de objetos documentales con objeto de identificarlos, describirlos, clasificarlos y localizarlos, pero los metadatos en la actualidad no se quedan ahí: pueden usarse como ya hemos mencionado también para indicar contexto, administración, procesamiento, preservación, disposiciones legales, etcétera, del recurso que se describe. Más aún, con los conceptos más actuales provenientes originalmente del SGML y en la actualidad del XML, pretendemos describir grandes y detalladas partes del contenido de un documento.

Los metadatos de la actualidad, a diferencia de los de épocas anteriores, pueden ser modificados o aumentados durante la vida del acervo, en función de que la biblioteca sea capaz de obtener y registrar información adicional de los documentos. Por tanto, en los acervos modernos, es común que los metadatos tiendan a incrementarse durante la vida de un documento. Los metadatos actuales no son ya creados solamente por las personas —catalogadores, usuarios, referencistas, etcétera— como antaño. Ahora pueden ser creados también automáticamente por una computadora, o “cosechados” o inferidos a través de relaciones con otro documento, tales como los hipervínculos.

Como puede apreciarse es de vital importancia que en las colecciones que pretendemos preservar para el futuro cuidemos el aspecto del registro y la organización documental de tales colecciones. Tan importante como la preocupación y cuidado que debemos tener por la resolución con que debemos digitalizar un documento o la duración del soporte en el que lo almacenaremos, es cuidar que el documento que se va a preservar cuente con los adecuados metadatos que permitirán recuperarlo en un futuro. Sobre todo si vamos a acumular numerosos documentos a lo largo del tiempo, metadatos adecuados nos permitirán encontrarlos después, a pesar de lo numerosos que sean. Aunque ello no resulta muy obvio para la mayoría de las personas –bibliotecarios incluso– los metadatos forman parte de la calidad intrínseca de un documento que se va a preservar, y por ello debemos cuidar, en primer lugar, que formen parte del documento, y en segundo lugar, que tengan una buena calidad. *Buenos metadatos contribuyen a tener buenos documentos* y, a la larga, a una *buena recuperación*. Y lo contrario también es cierto: es absolutamente imposible crear buenos documentos digitales con metadatos mediocres o nulos, y a largo plazo la recuperación de ellos será mínima o inexistente.

En cada proyecto que involucre acervos digitales debemos decidir cuál es el tipo de metadato que conviene a sus recursos y necesidades particulares. Y en general la respuesta de cuál tipo de metadato conviene usar se encuentra precisamente al balancear de manera correcta eso: los recursos y las necesidades. Hay que considerar que mejores metadatos insumen más recursos y viceversa. Por eso hablamos de un equilibrio entre ambos componentes. Y por lo mismo la respuesta de cuál tipo de metadato conviene usar en un proyecto dado variará en función de los documentos que se van a registrar, de la biblioteca, del acervo, del contexto y de los recursos que se tengan. Más importante aun que el tipo de metadato que se usará, será recordar que nunca deben construirse acervos digitales sin metadatos, y todavía más importante, que el tipo de metadato que se use para registrar el metadato tenga la mejor calidad posible y esté de acuerdo con los recursos disponibles.

Como ha podido observarse de todo lo anterior, los metadatos juegan un papel muy importante en la localización posterior de los documentos, sobre todo en colecciones de gran tamaño; esto incluye por supuesto a las colecciones de bibliotecas y archivos digitales. Y como ha podido verse también, los metadatos han tendido a la “especialización” y tratan en la actualidad de describir muy variados elementos del documento relacionado. En años recientes se ha desarrollado una preocupación muy particular para crear y perfeccionar una estructura relacionada con los metadatos más específicos para hacer una adecuada preservación documental.

La importancia de los metadatos para hacer preservaciones ha sido subrayada cada vez más por un gran número de organizaciones que han desarro-

lado algún tipo de esfuerzo en beneficio de sus actividades de preservación digital. Durante mucho tiempo fueron esfuerzos puntuales y aislados que carecían de coordinación interinstitucional o de estándares. Por eso observamos en años pasados la aparición de un gran número variantes de metadatos para preservación, los cuales obedecían a necesidades y contextos muy particulares de la organización que las creaba así como del tipo de documentos que querían preservar dentro de ellas. Pero en los últimos años se ha señalado cada vez más la necesidad de aprovechar la experiencia de esos esfuerzos aislados con miras a establecer consensos que permitan crear recomendaciones de tipo más general, así como estándares que puedan ser aprovechados por otras organizaciones y creen la infraestructura que necesita un sistema universal de metadatos específicos para hacer preservación.

Como conclusión final de este apartado podemos sostener que es de vital importancia para las colecciones que pretendemos preservar para futuras generaciones que cuidemos el aspecto del registro documental de las mismas; de otra manera serán poco utilizables. Tan importante como los aspectos técnicos y de calidad intrínseca del documento: resolución, integridad, escala, etcétera, es que nos preocupemos de cuidar que el documento que se quiere preservar cuente con los metadatos adecuados para que sus potenciales usuarios puedan recuperarlo eficientemente en un futuro, sobre todo ante la acumulación de numerosas colecciones a lo largo de años, décadas o siglos; no importa: adecuados metadatos nos permitirán encontrarlos después, a pesar de lo numerosos que lleguen a ser. Puede asegurarse también que en proyectos de preservación digital, la adición de buenos metadatos a los documentos implicará la adición de metadatos especializados para la preservación. Para ello ya hay proyectos al efecto que pueden servirnos de base y ser muy útiles en este sentido.

Como ha podido verse, los factores documentales son de suma importancia para la preservación digital. En palabras de Abby Smith:

La pregunta de fondo hacia las bibliotecas consiste en saber cuál es su ‘proposición de valor’ acerca de lo que ofrecerán en un futuro digital. A mi parecer será lo que siempre ha sido: su capacidad para explorar un inmenso universo de conocimiento que anda por ahí afuera, seleccionar un subconjunto de éste, ordenarlo con una descripción y acomodo coherente de tal forma que las personas puedan encontrar ahí fácilmente información confiable y auténtica. La única diferencia es que los bibliotecarios harán esa navegación en un universo mucho más grande.²¹

21 Abby Smith, Directora de Programas, Council of Library and Information Resources (CLIR), en: Wade, Roush, 2005, *The Infinite Library*, disponible agosto 2008 en: <http://www.technologyreview.com/Infotech/14408/>.

Con respecto a los otros tres factores enunciados, los culturales, sociales y económicos, han sido desarrollados en detalle en otro artículo. Para tener una idea de ellos los describo en forma sucinta:

Los factores culturales tienen que ver con la falta de sensibilidad de ciertas sociedades en ciertas épocas respecto al valor histórico de partes de su patrimonio documental, lo cual lleva a excluir la necesidad de guardar o archivar ese patrimonio y, al final, deriva en la pérdida de ese acervo documental. Nuestra época, contrariamente a lo que se pudiera pensar –se afirma que vivimos en la “sociedad de la información”– tiene en forma general un enorme problema de tipo cultural con respecto a la preservación de la información electrónica, una especie de oxímoron cultural.

Respecto a los factores sociales debemos pensar en cómo vamos a poder garantizar el acceso y la usabilidad documental. Es necesario establecer los mecanismos para que en un futuro se pueda tener acceso efectivo y masivo a esa información por parte del público. De nada sirve preservar por preservar si nadie tendrá acceso a ella. Salvados los derechos, propiedades y restricciones correspondientes, el reto futuro será cómo hacer llegar esa información digital a un número mayor de personas en el mundo sin crear nuevas brechas segregacionistas. Recuérdese todo lo que se ha debatido ya acerca de la “brecha digital” y sus consecuencias en el futuro. Debemos crear las infraestructuras para que sea posible que las personas tengan acceso a esa información y que la brecha tienda a cerrarse y no al contrario. Puede abundarse un poco más en este tema de la brecha en mi obra anterior *Un Modelo de bibliotecas Digitales para México*,²² y si se desea conocer más a fondo este problema, recomiendo la obra de Adolfo Rodríguez al respecto.²³

Finalmente un proyecto de estos alcances debe reflexionar insoslayablemente en materia de costos. ¿Cuánto cuesta hacer qué cosa? ¿quién tiene la responsabilidad de costear la preservación de acervos documentales? ¿cómo pueden lograrse economías? ¿existen modelos económicos sustentables? Éstos no son nunca proyectos baratos ni de corto plazo, por ello deben establecerse minuciosamente los elementos económicos que se verán involucrados en el desarrollo de esos proyectos para hacerlos y mantenerlos viables desde el punto de vista económico. Pero por caros que sean, debe recordarse siempre también a la hora de costear, que no preservar resulta a la larga todavía más caro.

22 Juan Voutsás M., 2007, *Un Modelo de Planeación de Bibliotecas Digitales Para México*, México, UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 309 p., ISBN: 970-32-4111-5. págs. 60-66

23 Adolfo Rodríguez G., 2006, *La Brecha Digital y Sus Determinantes*, México, UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 254 p., ISBN: 970-32-3853-X.

Los seis factores cubren tres enfoques fundamentales en este tipo de proyectos: el primero son los factores culturales y sociales que nos dan el enfoque humanista del proyecto, lo que nos ayuda mucho a ubicar los objetivos dentro del contexto social adecuado, la accesibilidad futura, los usuarios, la reducción de la brecha digital, etcétera. El segundo enfoque son los factores tecnológicos y documentales que nos ayudan a enfocar, por un lado los aspectos alrededor de la tecnología de cómputo y telecomunicaciones del proyecto, y por el otro la tecnología bibliotecológica y/o archivística alrededor del registro y recuperación documental. El tercer enfoque son los factores legales y económicos que nos permiten terminar de contextualizar el proyecto ayudándonos a precisarlo alrededor de estos aspectos jurídicos y financieros del proyecto, e incluir los intereses de editores y autores, el financiamiento, los costos, presupuestos actuales y proyectados para el futuro, etcétera.

CONCLUSIONES FINALES

Es esencial que existan los medios para identificar y mantener *para siempre* la información registrada. Sin ella, no habría registro del pasado, no habría manera de identificar, entender o replicar la forma de vivir de la sociedad en un momento dado –en este caso nuestra sociedad contemporánea–, ni los estudios y conocimientos recopilados en el pasado con objeto de hacer avanzar el conocimiento. Como ha podido verse hay una inmensa complejidad de tópicos que deberán ser considerados al emprender proyectos de preservación documental digital. El reto que plantean estas tareas no es fácil. Para aquellos que no conocían el problema de la digitalización documental y su preservación espero haber podido ilustrar un poco lo formidable de la tarea. Quienes ya habían pasado por este camino, espero que hayan podido aclarar algunas de sus dudas, ampliar sus horizontes, complementar su conocimiento y su experiencia al respecto, y con ello ayudarlos a realizar mejor su tarea. Como ha podido establecerse, el proceso de ofrecerles a comunidades de usuarios colecciones documentales digitales estructuradas y de preservarlas para futuras generaciones implica retos formidables que los bibliotecarios, archivistas y muchos otros profesionales de la información deben poder afrontar y resolver, y para ello deben conocer perfectamente esos retos a los que me he referido a lo largo de este documento.

Espero haber establecido con suficiente claridad que el proceso de digitalización documental y su preservación para uso posterior es un reto que implica amplios conocimientos tanto teóricos como metodológicos; decisiones delicadas, fuertes inversiones en tecnología y en recursos humanos calificados;

serias limitaciones legales, obsolescencia de dispositivos, reconversiones de tiempo en tiempo, ciertos riesgos, y el hecho de aceptar que el tratamiento del problema requiere de un enfoque multi y transdisciplinario. Muchos responsables de acervos han pospuesto su decisión de emprender proyectos acerca de preservación documental digital confiando en que llegaría un momento en que la tecnología y los costos se volverían estables y competitivos, sólo para darse cuenta de que a pesar de que la tecnología y sus costos mejoran cada día, la cantidad y variedad de materiales a digitalizar crece en forma exponencial, y de que este desvanece las ventajas ahora alcanzadas. Como espero haber establecido no es recomendable esperar a que los parámetros se estabilicen totalmente y los riesgos desaparezcan; eso no sucederá. Es necesario preservar ahora, y no hay otras alternativas que nos permitan darle vuelta o soslayar este camino tecnológico. Y como hay que recorrer ese camino, es mejor entonces entender el fenómeno ahora, prepararnos bien hoy para estos procesos, y conocer bien los riesgos y los vericuetos para evitar cometer errores haciéndolo bien desde el comienzo, sin incurrir en costos o riesgos innecesarios y sin cometer errores que a la larga debamos lamentar y/o corregir.

Es todavía común, sobre todo en nuestro medio mexicano, que muchos bibliotecarios y archivistas profesionales tiendan todavía a menospreciar el medio electrónico. Diversas razones —válidas y no— les permiten sólo voltear a ver la construcción de acervos documentales que están bajo soportes “tradicionales”. Bajo la premisa de “lo que no está impreso no está publicado”, simplemente tienden a ignorar el material digital en la red por considerarlo fútil, etéreo y riesgoso. Es cierto, la *Web*, como un todo, no es una biblioteca digital en sí misma, pero también es un hecho incuestionable que *todas las bibliotecas digitales forman parte de la Web*, así como una muy considerable parte de los archivos digitales. Y si todas las bibliotecas y esos archivos digitales forman parte de la *Web*, millones de documentos perfectamente válidos forman ya parte de acervos de bibliotecas y archivos, ahora electrónicos. Pero esos millones de documentos “formales” —y digo formales porque forman parte del acervo de esas colecciones digitales—, así como muchos millones más de documentos “informales” están en riesgo de perderse —y se pierden— con sobrecogedora frecuencia, precisamente por no querer, no saber o no poder hacer nada al respecto. Y si es el gran público el que no quiere, no sabe o no puede, ya sería grave, pero que esta situación se dé entre los profesionales de la información, es inadmisibile.

Hemos podido observar claramente cómo a nivel mundial la preocupación y los proyectos tendientes a estudiar y trabajar esta problemática han ido en aumento. Ahí están las grandes iniciativas norteamericana y la de la unión europea. Están también las nacionales de gran visión y alcance, como la australiana,

la canadiense y la holandesa, por citar algunas. En México, no contamos con ningún proyecto, programa, plan, iniciativa ni organización que actualmente trabaje en este rubro a gran escala. Todos los que existen son proyectos puntuales a nivel de ciertas instituciones puntuales. Reitero, nada estratégico ni de alcance nacional. El factor cultural nos afecta a todos gravemente para empezar. La falta de sensibilidad existe a todos los niveles. Las instituciones que tienen entre sus deberes la preservación de algún acervo documental, lo hacen principalmente con sus soportes “tradicionales”; muy rara vez o en muy poco porcentaje en lo relativo a acervos digitales, y por lo mismo, no han resuelto el problema; al contrario: hemos creado un enorme rezago como país en este sentido. Pero no debemos culpar a estas instituciones. Como hemos analizado la problemática a la que nos enfrentamos en este tipo de proyectos es formidable, no puede ser solucionada a nivel de organizaciones aisladas porque esto rebasa sus capacidades; debe ser enfrentada como nación, como región latinoamericana o, al menos, como grandes consorcios, con el concurso de múltiples sectores públicos y privados, además del académico. Lo mismo sucede en lo relativo a las personas que participen en ellos; este tipo de proyectos deben ser tratados con enfoques trans y multidisciplinarios además de que requieren el concurso de personas provenientes de muy variados sectores para lograr un adecuado diseño y puesta en marcha de un plan integral y armónico.

Uno de los principales propósitos de este artículo ha sido el de demostrar gradualmente a lo largo del mismo que la preservación digital no es un simple problema de tecnología y su obsolescencia; y que tampoco es sólo un problema que se solucione simplemente con convertir gradualmente todo lo que tenemos actualmente en papel en archivos digitales. Como hemos visto no se trata del simple almacenamiento de documentos sino de toda una escuela de pensamiento cuyo objetivo central pretende llevar la duración de materiales documentales hacia el largo plazo, muy largo plazo; idealmente a perpetuidad; que su hilo conductor implica la organización y observancia cuidadosa de procedimientos de gestión documental que involucran a todo el ciclo de vida de la información. En este sentido es muy importante concientizarnos de que la preservación digital no es sólo un método para mantener archivos de *bits* legibles y utilizables mañana, sino todo un proceso cultural y social que interactúa orquestada y armónicamente con una completa variedad de servicios que apoyen los ambientes de la información documental, y que contemplan, resuelvan e integren todos sus aspectos: técnicos, legales y económicos, todo dentro de un contexto de utilidad social. Al ser entonces un proceso cultural y social, deberá involucrar a diversos actores de variados sectores de la sociedad: autores, editores, bibliotecarios, archivistas, gobernantes, legisladores, informáticos, planeadores, tomadores de decisiones, etcétera, y se deberá repartir entre

ellos armónica y equitativamente la responsabilidad que a cada uno compete, en especial la responsabilidad social. A pesar de la dificultad, los riesgos y los costos de preservar, es imposible soslayar el problema. Todos los conglomerados humanos, todas las regiones y países poseen riquezas documentales que necesariamente hay que preservar y distribuir porque forman en esencia su patrimonio cultural; México no es la excepción. Y nuestro patrimonio cultural y documental es enorme, rico y variado; es impensable no considerarlo y no tratar de preservarlo.

Los aspectos sociales de la preservación documental deben ser analizados siempre junto con los demás aspectos para tener una visión completa de ella. Estos proyectos no pueden ser vistos sólo desde enfoques administrativos, tecnológicos o legales. Habrá que agregarles siempre el enfoque de las personas y su contexto social; deberemos poder garantizar el acceso y la usabilidad documental para la generación actual y las futuras; será imperativo establecer los principios, valores y premisas para que en un futuro el público pueda tener acceso efectivo y masivo a esa información, y tendremos que velar siempre porque así suceda, sobre todo entre profesionales de la información. Hay un principio social básico que rige este aspecto: “de nada sirve preservar por preservar la información si nadie, –o muy pocos– tendrán acceso a ella”. Salvados los aspectos técnicos, de propiedad intelectual y financieros que hay que cubrir para que la información se preserve, el reto máximo para el futuro será cómo hacer llegar esa información digital a un número cada vez mayor de personas en el mundo sin crear nuevas brechas segregacionistas. Como establecimos, no deseamos que la preservación digital sea uno más de los elementos que amplíen la “brecha digital” y sus consecuencias. Debemos pensar ahora en cómo crear las infraestructuras para garantizar que en el futuro las personas tendrán acceso a esa información y que la brecha tienda a cerrarse.

REFERENCIAS BIBLIOGRÁFICAS

* *Todas las referencias electrónicas han sido verificadas como existentes y exactas hasta el 1 de agosto del 2008.*

Beagrie, Neil. 2006. “Digital Curation for Science, Digital Libraries, and Individuals”, en *The International Journal of Digital Curation*, Volumen 1, núm. 1, Otoño 2006, disponible agosto 1, 2008 en: <http://www.ijdc.net/ijdc/article/viewFile/6/49>

Besek, June M., 2005. *Copyright Issues Relevant to Digital Preservation and Dissemination of Pre-1972 Commercial Sound Recordings by Libraries and Archives*, (Washington, D.C.: Council on Library and Information Resources and the Library of Congress, Dec. 2005), p. 2.

- Brooks, Tim, 2005, *Survey of Reissues of U.S. Recordings* (Washington, D.C.: Council on Library and Information Resources and the Library of Congress, Aug. 2005).
- Buckner Higginbotham, Barbara. 1990, "The 'Brittle Books Problem': A Turn-of-the-Century Perspective", en *Libraries and Culture*, n. 25, v. 4 (Fall 1990): 496-497.
- Committee on Preservation and Access, 1986, *Brittle Books*, Reports of the Committee on Preservation and Access (Washington, D.C.: Council on Library Resources, 1986), p. 21;
- "Copyright Law of the United States of America and Related Laws Contained in Title 17 of the *United States Code*", disponible agosto 1, 2008 en: <http://www.copyright.gov/title17/92chap1.html>.
- Conway, Paul, 1996, "Preservation in the Digital World", en *Council of Library and Information Resources (CLIR) Publications*. Disponible agosto 1, 2008, en: <http://www.clir.org/pubs/reports/conway2/>.
- DLF -Digital Libraries Federation-, 2002, *Benchmark for Faithful Digital Reproductions of Monographs and Serials*, The Digital Library Federation Benchmark Working Group (2001-2002), Página Oficial de la DLF, disponible agosto 1, 2008 en: <http://www.diglib.org/standards/bmarkfin.htm>
- Duranti, Luciana, and Thibodeau, Kenneth, 2005, "The concept of record in interactive, experiential and dynamic environments: The view of InterPARES", en *Archival Science*, Springer Netherlands, ISSN:1389-0166, (Print) 1573-7519 (Online), Vol. 5 Núms. 2-4, December 2005., DOI 10.1007/BF02660804, pp. 13-68.
- Gavrel, Sue, 1986, "Preserving Machine-Readable Archival Records : A Reply to John Mallison", en *Archivaria* 22(Summer),153,55.
- Gladney, Henry M. 2004, "Preserving Digital Records : A Method Guided by Scientific Philosophy", en *ERPANET*, Electronic Resource Preservation and Access Network, Página de la Red, disponible agosto 1, 2008 en: http://eprints.erpanet.org/71/01/TDO_4_Archivaria_submitted_17_Nov.pdf.
- Gorman, Michael, 1998, "What is the Internet", en *The One Person Library*, vol. 15, Núm. 6 (6/98) p. 5.
- Herring, Mark Y. 2001, "10 Reasons Why the Internet is no Substitute for a Library", en *American Libraries*, April 2001, pp. 76-78, disponible también agosto 1, 2008 en: "American Libraries OnLine : News and Updates from the Magazine of the American Library Association", Página Oficial de la ALA: <http://www.ala.org/ala/online/selectedarticles/10reasonswhy.cfm>.
- IFLA International Federation of Library Associations, 2006, *Manifiesto de la IFLA/UNESCO Sobre Internet : Directrices*, compilado y Editado por Seidelin, Susanne, Directora de la Oficina IFLA/FAIFE, Septiembre 2006, Página Oficial de la IFLA, disponible agosto 1, 2008 en: <http://www.ifla.org/faife/policy/iflstat/Internet-ManifiestoGuidelines-es.pdf>.

- Instituto de Investigaciones Jurídicas, Universidad Nacional Autónoma de México (UNAM), 2000, "Ley Federal del Derecho de Autor", disponible agosto 1, 2008 en: <http://info4.juridicas.unam.mx/ijure/fed/139/default.htm?s=>.
- Lyman, Peter y Varian, Hal. 2000, "How Much Information?", en *Journal of Electronic Publishing*, Diciembre, 2000, vol. 6:2. ISSN 1080-2711, disponible agosto 1, 2008 en: <http://www.press.umich.edu/jep/06-02/lyman.html>
- _____, 2003, "How Much Information? 2003", en Sitio Web Oficial de la Universidad de California en Berkeley, EUA, disponible agosto 1, 2008 en: <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>.
- Lynch, Clifford, 1997, "Searching the Internet", en *Scientific American*, March 1997, disponible agosto 1, 2008 en: <http://www.hackvan.com/pub/stig/articles/trusted-systems/0397lynch.html>.
- Mallison, John C. 1986, "Preserving Machine-Readable Archival <http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/> Records for the Millenia", *Archivaria* 22 (Summer): 147,52;
- Rodríguez G., Adolfo, 2006, "La Brecha Digital y Sus Determinantes", México: UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 254 p. ISBN: 970-32-3853-X.
- Russell Young, John, 1897, *Report of the Librarian of Congress*, December 9, 1897 (Washington: GPO, 1897), pp. 49-50.
- Smith, Abby, Directora de Programas, Council of Library and Information Resources (CLIR), en: *Wade, Roush*, 2005. *The Infinite Library*, disponible agosto 1, 2008 en: <http://www.technologyreview.com/Infotech/14408/>.
- Soules, Aline, 2002, "The Rights and Wrongs of the Doctrine of the First Sale", en *Electronic Book Web*, disponible agosto 1, 2008 en: [http://12.108.175.91/ebookweb/stories/storyReader\\$1816](http://12.108.175.91/ebookweb/stories/storyReader$1816) .
- Voutssás M., Juan, 2006, *Bibliotecas y Publicaciones Digitales*, México : UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 342 p. ISBN:970-32-3962-5.
- _____, 2007, *Un Modelo de Planeación de Bibliotecas Digitales Para México*, México : UNAM, Centro Universitario de Investigaciones Bibliotecológicas, 309 p. ISBN: 970-32-4111-5.

